Biblioteka Informatyki
Szkół Wyższych

Information Systems Architecture and Technology

The Use of IT Technologies to Support Organizational Management in Risky Environment

Library of Informatics of University Level Schools

Series of editions under the auspices
of the Ministry of Science and Higher Education

The ISAT series is devoted to the publication of original research books in the areas of contemporary computer and management sciences. Its aim is to show research progress and efficiently disseminate current results in these fields in a commonly edited printed form. The topical scope of ISAT spans the wide spectrum of informatics and management systems problems from fundamental theoretical topics to the fresh and new coming issues and applications introducing future research and development challenges.

The Library is a sequel to the series of books including Multidisciplinary Digital Systems, Techniques and Methods of Distributed Data Processing, as well as Problems of Designing, Implementation and Exploitation of Data Bases from 1986 to 1990.

Wrocław University of Technology

Information Systems Architecture and Technology

The Use of IT Technologies to Support Organizational Management in Risky Environment

Editors
Zofia Wilimowska
Leszek Borzemski
Adam Grzech
Jerzy Świątek

Wrocław 2014

Publication partly supported by
Faculty of Computer Science and Management
Wrocław University of Technology

Project editor
Arkadiusz GÓRSKI

The book has been printed in the camera ready form

OFICYNA WYDAWNICZA POLITECHNIKI WROCŁAWSKIEJ
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław
http://www.oficwyd.pwr.wroc.pl;
e-mail: oficwyd@pwr.wroc.pl
zamawianie.ksiazek@pwr.wroc.pl

ISBN 978-83-7493-858-7

CONTENTS

Introduction 5

PART 1. MODELS OF INFORMATION MANAGEMENT FOR ORGANIZATIONS

1. Cezary ORŁOWSKI, Arkadiusz SARZYŃSKI
   A Model for Forecasting PM10 Levels with the Use of Artificial Neural Networks . 15
2. Norbert BOLEŃSKI, Edward SZCZERBICKI
   Knowledge-Based Management as the Key Success Factor for Research and Development Organizations 31
3. Maria GALANT-PATER, Anna LAMEK, Maciej MALARA, Bogusz PRZYBYSŁAWSKI, Witold REKUĆ, Leopold SZCZUROWSKI
   Ontology as a Tool for Modelling Business Processes of Transport 41
4. Ryszard ANTKIEWICZ, Michał DYK, Rafał KASPRZYK, Andrzej NAJGEBAUER, Dariusz PIERZCHAŁA, Zbigniew TARAPATA
   Modelling, Simulation and Computer Support of the Polish Criminal Procedure 51
5. Dorota KUCHTA
   Machine Failure Oriented Robust Schedule in Flow Shop Scheduling – a Fuzzy Approach 61
6. Jarosław A. CHUDZIAK, Andrzej CIEMSKI
   Acquisition, Generation and Management of the Customer Knowledge in an Interactive Marketing 71
7. Aleksander ORŁOWSKI, Edward SZCZERBICKI
   The Case of Management Concepts in Civil Aviation Business 81
8. Paweł KUŻDOWICZ
   Modeling the Accounting Systems in Standard Software 91
9. Zofia WILIMOWSKA, Agnieszka PARKITNA, Arkadiusz GÓRSKI, Anna Maria KAMIŃSKA, Anna KIŁYK, Katarzyna GWÓŹDŹ
   Assessment of the Financial Condition of SMES: is There a Need to Establish New Compulsory or Voluntary Financial Standards? 103
10. Mariusz FRAŚ, Ziemowit NOWAK
    On Some Issues of Web Content Processing Efficiency on the End User Side 113

PART 2. RISK MANAGEMENT IN DECISION MAKING PROCESS

11. Bogdan RĘBIASZ, Bartłomiej GAWEŁ, Iwona SKALNA
Capital Budgeting of Interdependent Projects with Fuzziness and Randomness 125
12. Adriana HALIKOWSKA, Daniel HALIKOWSKI
Attractiveness Study of Specific Industry in Opole Voivodship and Poland 137
13. Michal J. KOWALSKI
Value Creation Measures as a Support in Investment Decisions Concerning Merges and Acquisitions 147
14. Danuta SERETNA –SAŁAMAJ, Agnieszka SZPARA
Adjustment of the Financial Company Structure in the Process of Investment, Including the Risk Sensitivity Coefficient 159
15. Maciej ZARĘBA
Application of 2-Additive Measure to Optimization of Diversification Across Sectors in Portfolio Selection Problem 173
16. Katarzyna ROSTEK , Agnieszka SKALA
In Search of High-Tech Entrepreneurship: an Algorithm For Differentiation of Business Entities Operating within the High-Tech Sector as Illustrated by the Example of Warsaw 185

INTRODUCTION

THE USE OF IT TECHNOLOGIES TO SUPPORT ORGANIZATIONAL MANAGEMENT IN RISKY ENVIRONMENT

Risk is an attribute of all human activities, in particular business. It occurs always. Uncertainty is a feature of reality, indicating the impossibility of accurately determining future events. Its sources are objective factors arising from the variability and complexity of the situation and the subjective, involving people with mental processes. Risk management, in accordance with the risk management standard published by the Federation of European Risk Managers Association should be a continuous and developing process which includes both the organization's strategy and the implementation of that strategy. Should address methodically all the risks surrounding the organization's activities, which took place in the past, present and the greatest – may occur in the future.

The company operates in an environment that affects its actions and, consequently, to assess its value. Unpredictable external and internal changes of the organization and the ongoing relationship between them mean that a potential investor feels greater uncertainty. In the literature, are considered two basic approaches to risk analysis in business management:

• draw attention to the effects of the presence of risk;

• draw attention to the cause of the risk .

These two approaches imply how to integrate risk management decisions, as well as different ways to reduce risk. In the literature, one can distinguish the two approaches to risk measurement:

• measurement of one type of risk; identification of specific types of risk, its measurement and the application of appropriate mitigation techniques.

• modeling decision problems, taking into account risks; application of methods and techniques of modeling of economic phenomena taking into account the risks (eg. stochastic models, fuzzy models, computer simulation models etc.) and making the best decision in a particular decision situation.

Sources of uncertainty associated with managerial decisions can be external (relate to the environment of the company) and internal (relate to the managers' decisions)

Variability and complexity of the environment mean that the risk is considered to be a common phenomenon associated with almost every activity. It should not be

considered only as a threat, because often creates opportunities inspirational human actions. For this reason, the risk should be considered as one of the determinants of entrepreneurship. It has helped to conduct numerous studies, the creation of new theories and attempts to apply them in practice. It can therefore be regarded as a source of progress and activation, and therefore considered the risk to be useful phenomenon. It will be so only if you give up control, or subjected to inspection and control.

The risk is defined in many ways and can be:

- the possibility of incurring losses
- the probability of loss,
- discrepancy between actual and expected results,
- probability of a different result than expected,
- the potential for adverse events,
- conditions in which there is a possibility of loss,
- uncertainty, danger, to realizing purpose, expected outcomes established at the time of making an investment decision.

The main source of risk in human activities is uncertainty as to the degree of attainment of the objective in the future, resulting from the fact that decisions are made today, and the effects of the decision are known only in the future. Uncertainty is a feature of reality, indicating the impossibility of accurately determining future events. Its sources are objective factors arising from the variability and complexity of the situation and the subjective, involving the mental processes decision makers.

An improvement of decision making process is possible to be assured by analytical process supporting. Applying some analytical techniques, such as: computer simulation, expert systems, genetic algorithms can improve quality of managerial information. Managers have to remember that "What's worked before, especially in the recent past, becomes an acceptable pattern of behaviour – except that the market will then do something similar, but different enough to deceive us" Combining analytical techniques and building computer hybrids give synergic effects – additional functionality – which makes managerial decision process better. Different technologies can help in accomplishing managerial decision process, but no one is in favour of information technologies, which offer differentiable advantages.

The computers are able to collect and select the information can make some statistics, but decisions have to make managers basing on their experience and taking into consideration computer support.

The book entitled *Information Systems Architecture And Technology – The Use of it Technologies to Support Organizational Management in Risky Environment* focuses on very hot topics in the field of today's various computer systems based applications and is devoted to information systems concepts and applications supporting exchange of goods and services by using different business models and seizing opportunities offered by IT systems.

The selected and reviewed chapters have been divided into two parts:
PART 1. **MODELS OF INFORMATION MANAGEMENT FOR ORGANIZATIONS**
PART 2. **RISK MANAGEMENT IN DECISION MAKING PROCESS**

The book provides an interesting representation of research in the area of contemporary management information systems - concepts, models, services and applications, requirements and models.

The first part – **MODELS OF INFORMATION MANAGEMENT for ORGANIZATIONS** – presents considerations of information based organizations. Here is considered process of collecting, representing, protecting and distributing the business information. Here are showed examples of models and practical use that information in managing process. Dynamic development of the IT technologies creates the possibility of using them in the dynamic management process modeling and decision making processes supporting.

The second part **RISK MANAGEMENT IN DECISION MAKING PROCESS** considers special models designed for finance and investment management. It starts with budgeting model and then discusses some special models for financial and investment decisions taking focusing on various methods that are risk undertaking in managerial decision process.

PART 1. MODELS OF INFORMATION MANAGEMENT FOR ORGANIZATIONS

Dynamic development of the IT technologies creates the possibility of using them in the dynamic management process modeling and decision making process supporting.

The **Chapter 1** describes a method of forecasting the level of PM10 with the use of artificial neural networks. The World Health Organization estimates that over two million people die every year as a result of complications caused by PM10. Organizations around the world seek to limit its emissions. Meteorological data was taken into account in the construction of the model, and unidirectional networks were used to implement it due to their ease of learning. The obtained results were then compared with real values and the complete network configuration (minimizing the forecast error) was determined. After completion of the learning process, the developed network was used to forecast the particulate matter levels in Gdansk.

The aim of the **Chapter 2** is to outline the area of knowledge management within this type of organizations and to describe the model of management. Project organizations are organizations in which each task is carried out individually. They are characterized by a lack of repeatability due to the necessity of matching specific solutions to specific customer. A key to achieving success in project organizations lies in exemplary management of processes within those organizations, while the ongoing projects are mainly characterized by their uniqueness.

The **Chapter 3** is directly associated with the task that consists in developing a methodology for identifying and analyzing the problems of transport services provision, and finding solutions to these problems by using a specific IT platform. Modern approaches to solving problems associated with provision of transport services should be based on the use of widely understood knowledge of the subject domain. The subject of this study is an analysis of the solutions in the scope of transportation ontology available in literature. The review of literature carried out as a part of this study shows that with respect to transportation ontology the offer is quite wide and largely covers the area of interest of the authors.

Current analysis and social research on the Polish judiciary indicate negative opinion stated by approximately 40% of the Polish citizens. In particular, the high costs and long delays of criminal cases are criticized. In the **Chapter 4** the Authors show that it is possible to treat the criminal procedure in terms of formal models (e.g. business processes or stochastic processes) and there is presented IT system supporting tasks of a criminal department of a district court in the two dimensions: (1) processing of information, flow documents and work scheduling; (2) modelling, simulation and optimization of criminal procedure. For the determination of the quantitative characteristics of the criminal procedure's actions the Authors propose to apply models of queuing networks, models of "Activity Network" and "Critical Path" – all of them supported by a dynamic discrete-event simulation.

In the **Chapter 5** a new concept of the robustness of flow machine schedule is proposed, which is based on the features of individual machines. Its aim is to protect the organization against serious delays caused by accumulation of delays in various processes realized in the organization at the same time and due to machine failures. A fuzzy model determining the robust schedule defined in the paper is formulated, which can be reduced to a mixed integer linear model. The approach is illustrated by means of an example.

The **Chapter 6** presents an approach and experiences in building business and technology environment for acquisition, generation and management of the customer knowledge in interactive marketing and shows a customer insight process from acquisition up to analytical models which provide proper, treatment-oriented segmentation with propensity models. It describes also use of mathematical models and other necessary tools to build analytical solution for continuously increasing demand for adequate data from interactive marketing.

The **Chapter 7** refers to the papers presented by the authors on the previous conferences. Previously the area of interest was concentrated on the process of creation of models dedicated to help in decision making processes in the area of managing airlines and airports. In this chapter are presented different concepts of management which were and are in use by international airlines. Each concept is going to be presented with strengths and weaknesses of it to get the analyze of current market concepts which are used by international airlines. The main changes in management concepts are presented, basing on the example of one airline, LOT Polish Airlines.

Changes which take place in finance sector starting from the information which is required by the stakeholders and ending at dynamic IT development forced most of the accounting systems evolution. The **Chapter 8** shows some aspects of the modeling system based on the ERP systems (Enterprise Resource Planning). Presented method focus on selected value steam flows (for example production are, sales) including the accounting records on general ledger accounts and on costs and achievements account objects.

The **Chapter 9** is an discussion about the problems of the financial condition assessment standards in the SME sector in Poland. Herein is conducted an analysis and diagnosis of the state of the existing literature and business practice. The article suggests that there is a need to adapt to the demand of the practice of literary activity by identifying shortcomings. The result is a plan, which includes the most important steps which are necessary to be taken in research methodology to solve current problems of SME sector.

The **Chapter 10** tackles the topic measuring Web services efficiency with emphasis on client side processing. A brief description of existing approaches to Web service quality analysis is followed by an explanation why client side processing is gaining in importance. Analysis of HTML documents processing and the navigation timing API is described in order lay the foundations of the experiment procedure and measurements. Chapter is concluded with a presentation of results based on a comparison of HTML5 and Flash in popular Web browsers followed by an analysis of the findings.

PART 2 RISK MANAGEMENT IN DECISION MAKING PROCESS

The main objective of all management decisions taken is the company's growth benefits to the owners which can be realized by maximizing the market value of the company. It is a difficult process because it affects all business areas and because the risk which is associated with business activity influences by external and internal changes.

Capital budgeting problem is concerned with allocation of a company’s capital to a suitable combination of projects such that the investment can bring the company the maximal profit. The **Chapter 11** discusses the hybrid investment projects portfolio selection problem in the situation where only some parameters are well described by random variables and the some parameters can be hardly predicted by historical data and can be described by fuzzy variable. The model takes into account stochastic and economic dependency between projects. A hybrid solution method of this model produces a Pareto optimal set. As an illustration, an example taken from metallurgical industry is also provided.

In the **Chapter 12** there is discussed industry attractiveness in the region as the factor that affect the success of investment is. Often one of the ways of investing is opening new branches, that are supposed to become an important pillar of the whole

enterprise. The Authors used data published by Central Statistical Office of Poland to determine relation which allows to answer to the question whether chosen by the company industry is competitive. For the calculations authors used principal component method based on eigenvectors and eigenvalues. On this method basis the degree of explanation of attractiveness for Poland, Opole Voivodeship and for the concentration coefficient was determined

Mergers and acquisitions are, beside the organic growth, one of the main paths to increase capital value. They are currently presented as a natural part of corporate practice, and as such appear in strategy textbooks and MBA classes. In the **Chapter 13** a proposal of operating procedure in acquisition decision-making was presented. A special attention was paid to transactions where only part of shares are acquired. A detailed propositions of value measures adjustments concerning acquisition accounting were also included. A detailed solutions for goodwill, badwill and restructuring costs were proposed. An example of the proposed approach was given.

Management of financial structure of the company can be considered as one of the most difficult issues in the whole process of business management. The **Chapter 14** presents the dynamic development of the financial structure in the process of investing with regard to sensitivity to risk. Uncertainty surrounding economic activity creates the need for a dynamic view of the structure of finance in a changing environment and especially in the process of investing. Investment decisions inherent risks. Decisions regarding the management structure of finance in a company have a significant impact on the financial position and consequently the value of the company.

The **Chapter 15** shows a multistage method of portfolio selection based on: the risk-return, book value to market value ratio and Choquet integral. Proposed model uses data from stock price listed on the Warsaw Stock Exchange (at least since 2000) and are part of selected sector index. Obtained results were compared to the results of APT portfolio and market portfolio (WIG). An example of the above approach was given.

Manufacturing companies operating within the high-technology sector (HT) are of interest to science, industry and national authorities because of the characteristics ascribed to them. From the macroeconomic perspective, the sector is important for several reasons. So it is important to identify the business entities belonging to that sector. The **Chapter 16** presents an algorithm for differentiation of business entities operating within the high-tech sector. To identify the entities belonging to the high-technology industry, it is necessary to perform a sequence of activities which form the procedural algorithm. Usefulness of the algorithm has been verified using the example of a group of Warsaw high-tech companies which were subject to investigation under the European project entitled: “Stołeczne Forum Przedsiębiorczości” (Warsaw Entrepreneurship Forum). In the future, the algorithm could be used as the basis for the implementation of an IT tool for the identification and description of high-tech businesses.

Researchers should note that the model is only a representation of reality, is only a simplified representation of a given segment of reality, which contains a certain number of its properties relevant to the research. The results of the model, even using the best computer tools, should be evaluated by humans.

The book contains contributions accepted by reviewers of the submitted works. We hope that the book will be considered as a forum for presentation of original works on a good professional level and for discussions integrating different subjects of enterprise management and changes as well as information systems planning, designing, development and implementation.

We thank all the Authors who have submitted their works to be published in this book and the Reviewers that they were so kind to prepare the reviews.

Wroclaw, September 2014

Zofia Wilimowska

PART 1

MODELS OF INFORMATION MANAGEMENT FOR ORGANIZATIONS

Forecasting PM10, artificial neural network

Cezary ORŁOWSKI*
Arkadiusz SARZYŃSKI

A MODEL FOR FORECASTING PM10 LEVELS WITH THE USE OF ARTIFICIAL NEURAL NETWORKS

This work presents a method of forecasting the level of PM10 with the use of artificial neural networks. Current level of particulate matter and meteorological data was taken into account in the construction of the model (checked the correlation of each variable and the future level of PM10), and unidirectional networks were used to implement it due to their ease of learning. Then, the configuration of the network (built on the basis of the developed model) was established, defining the number of layers and neurons, as well as the activation function. 4 methods of propagation (Back Propagation, Resilient Propagation, Manhattan Propagation and Scaled Conjugate Gradient) were applied in the network learning process to select the best method. The obtained results were then compared with real values and the complete network configuration (minimizing the forecast error) was determined. After completion of the learning process, the developed network was used to forecast the particulate matter levels in Gdansk.

1. INTRODUCTION

The consequence of technological and industrial development in the modern economy is the increase in emissions of harmful substances and particulates. PM10 is especially dangerous for living organisms. The World Health Organization estimates that over two million people die every year as a result of complications caused by PM10. Organizations around the world seek to limit its emissions. The maximum allowable average daily level of particulates is 50μg/m^3 and it is necessary to control this level. What is also important is that the level needs to be forecast, as operational activities can be determined on the basis of such a forecast, especially in the case when the alarm level is exceeded. The existing forecasting methods only allow for the

* Gdańsk University Of Technology, Faculty of Management and Economics, Department of Information Technology Management

determination of the average daily level. This paper will present an attempt to build a model for forecasting hourly levels of PM10 (24 hours into the future). The model will then be implemented using a unidirectional network. Thus, conditions for forecasting the level of PM10 for 24 hours in advance will be created, which will have a reasonable limit of error.

2. THE CONSTRUCTION OF THE MODEL TO PREDICT THE LEVEL OF PARTICULATE MATTER

The first stage necessary to begin the work was the construction of the model. The input and output variables were defined after initially identifying the variables which affect the level of PM10. It was established that the future value of the PM10 level depends on the prevailing meteorological conditions and the current value of the level of particulate matter [2]. It was also concluded that virtually all meteorological data may affect the future level of PM10. Therefore, each of them was studied independently to select only those parameters that have an impact on the level of air pollution. Research was carried out with the use of the method of statistical analysis to examine the correlation coefficients of the individual parameters.

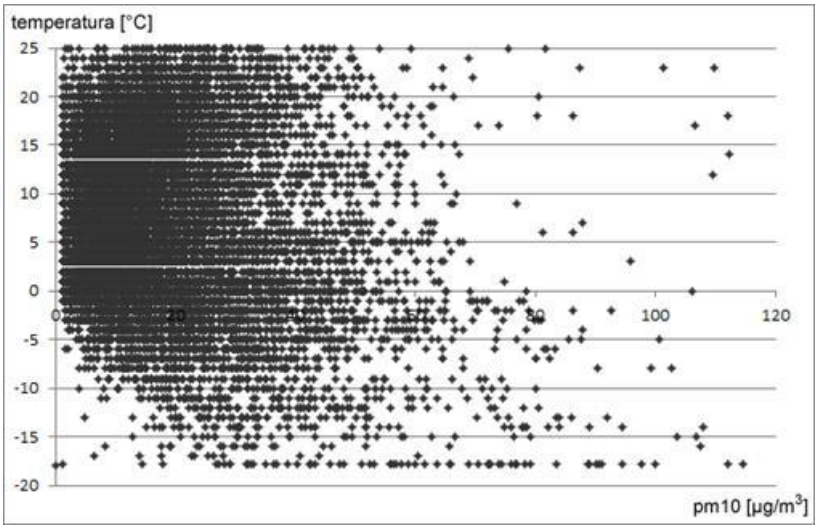

Fig. 1. The results of the correlation analysis of variables: air temperature and PM10

Figure 1 shows the results of the correlation analysis of variables: air temperature and PM10. Data used in the study involved temperature values ranging from -20 to 25 degrees and the level of particulate matter from 0 to 120 $\mu g/m^3$. It can be concluded (on the basis of the data in Figure 1) that the level of particulate matter decreases with

increasing temperature. Thus, changing air temperature can be used in the construction of the model.

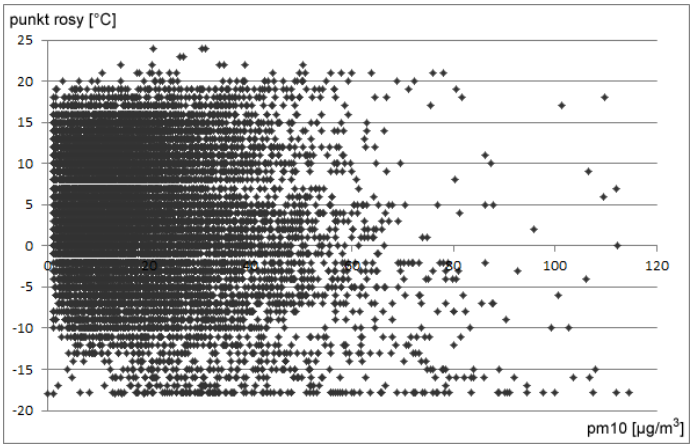

Fig. 2. The results of the correlation analysis of variables: dew point and PM10

Another analysis of a variable, from the perspective of the model proposed in this work, is the analysis of the impact of the dew point (in the range from -20 to 25) on the level of PM10. Figure 2 (the correlation of the dew point and the level of PM10) shows that the level of particulate matter decreases with the increasing dew point. For this reason, the dew-point variable can be used in building the model.

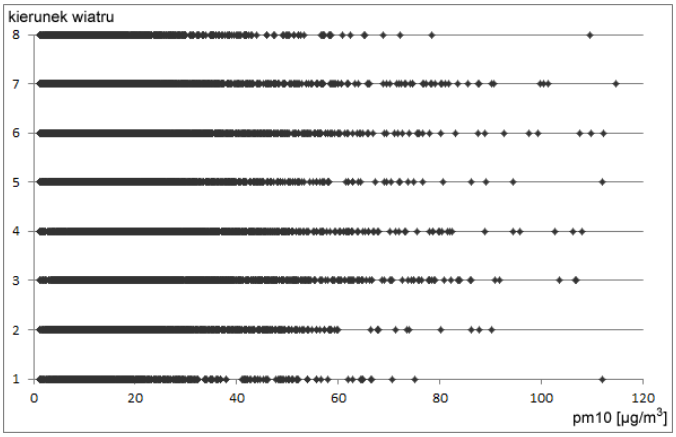

Fig. 3. The results of the correlation analysis of variables: wind direction and PM10.

Another variable analyzed in reference to the proposed model was the wind direction. The values of this variable (0-360 degrees) were assigned (for ease of calculation) to one of the eight zones (0-45, 45-90, 90-135, 135-180, 180-225, 225-270, 270-

315, 315-360). The correlations of the variables: wind direction and the level of PM10 were analyzed. Figure 3 shows the results of these analyses. A relationship between these variables was identified. The variables corresponding to the adopted zones have an impact on the level of particulate matter. For example, the wind blowing from the direction marked as 1 and 8 caused much less indication of the level of PM than that blowing in the direction marked as 3 and 7.

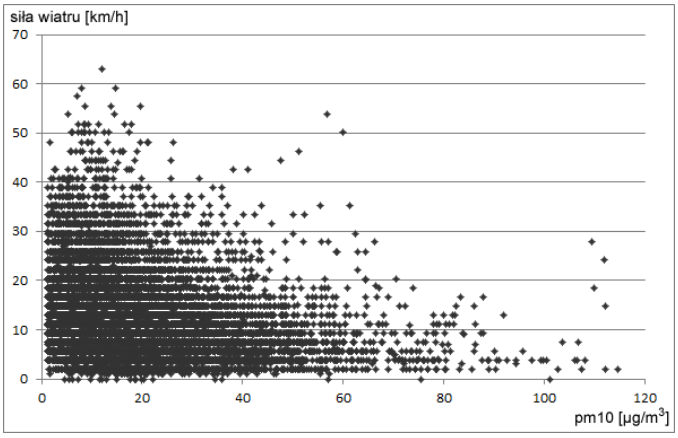

Fig. 4. The results of the correlation analysis of variables: wind power and PM10.

The power of the wind was another analyzed variable. Analyses similar to those presented above were carried out. According to Figure 4 (showing the correlation of the wind power variable and the level of particulate matter variable), this correlation is high. In low wind, the level of PM10 reaches a higher level, and when the wind is strong, the content of PM in the air is lower.

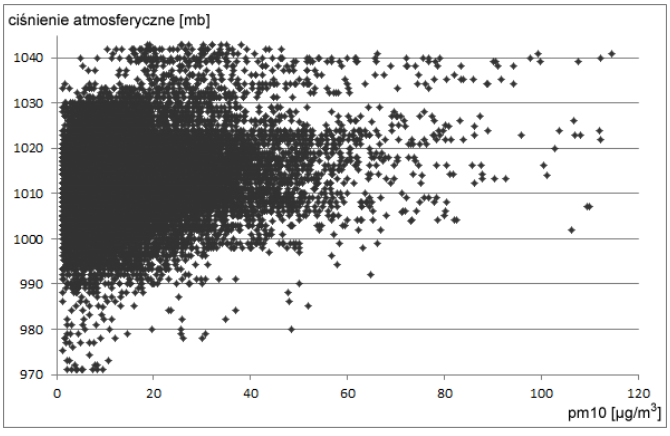

Fig. 5. The results of the correlation analysis of variables: air pressure and PM10

Atmospheric pressure is another variable which needs to be taken into account in the construction of the model. Figure 5 shows that the level of correlation of the variables: air pressure and the level of particulate matter is significant. It is visible that an increase in the level of PM10 takes place with an increase in atmospheric pressure.

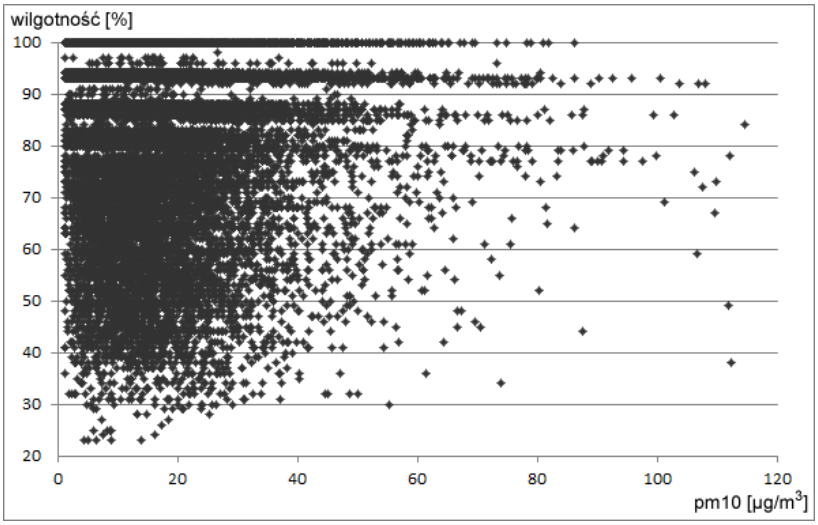

Fig. 6. The results of the correlation analysis of variables: humidity and PM10.

The last of the studied variables is humidity. When analyzing the correlation of the variables: the level of air humidity and PM10 (Figure 6) it was discovered that, as in the case of atmospheric pressure, the level of air pollution increases with increasing humidity. Similarly (to the atmospheric pressure variable) this variable also affects the level of PM10.

After conducting the correlation analyses, the results of which are shown in Figures (1-6), the structure of the input variables in the proposed model was established. The input variables are to include the current value of the level of air pollution with PM10, as well as 6 essential weather parameters (air temperature, dew point, wind direction, wind power, atmospheric pressure and humidity). A change in any of these parameters differently affects the future value of PM10 levels, either raising or lowering its value. It is thus necessary to take into account all of these variables in predicting the level of air pollution.

The next step in the construction of the model was to specify the output variables. The aim of the study was to forecast the future value of PM10 levels, therefore the output variable achieved after applying the model should be a single variable, corresponding to the future value of the level of air pollution with PM10. To sum up, the proposed model consists of 7 input variables.

3. THE IMPLEMENTATION OF THE MODEL – UNIDIRECTIONAL MULTI-LAYER ARTIFICIAL NEURAL NETWORK

The following stage was the implementation of the model described above. The method of implementation had to be chosen on the basis of the model, which would allow the forecast of the future value of the PM10 level. The issue of forecasting particulate matter is a non-linear problem. An artificial neural network is a method which is used for solving this type of problem. It is a method involving modelling complex non-linear functions with numerous independent variables [3]. In the case of forecasting PM10 levels, the independent variables are meteorological factors and the current level of PM10. There are no predetermined algorithms processing data within the network. It is assumed that the network adapts itself to the problem being solved. This is the main advantage of neural networks - their versatility and applicability in terms of solving various issues, including predicting the future. The choice of a neural network as the forecast method is therefore justified.

Two main types of neural networks were taken into consideration: unidirectional and recursive. For most issues, unidirectional networks are implemented. This is due to the fact that they assist in solving the majority of problems, and are also much easier to implement. Recursive networks are used to solve more complicated problems, such as complex optimization problems. Hence, a unidirectional network was chosen for the forecasting of the PM10 level, and also used for solving other prediction problems.

An important step in the construction of a unidirectional neural network is to determine whether a single-layer or multi-layer network will be used. Single-layer networks are of little use. Apart from the input layer (where no calculations are made), it only has one layer of neurons (individual network elements), which serves as the output layer. The neurons of this layer operate independently of each other. The capability of such a network is limited to the capability of individual neurons. Single-layer networks are used mainly in simple decision problems. In the case of forecasting, a multi-layer network needs to be used. It is characteristic for multi-layer networks that in addition to the input and output layer, they have at least one hidden layer. Through the use of more than one layer, they allow closer mapping of any given complex issue. The vast majority of problems can be solved with the use of only one hidden layer, including problems associated with forecasting. More complicated problems, such as image recognition, require the use of two or more hidden layers. The constructed network will contain only one hidden layer.

The next step in the construction of the neural network is to determine the number of neurons in each layer of the network (the input, output and hidden layers). The input layer contains as many neurons as there are input variables. The adopted model assumes 7 input variables so there will be exactly 7 neurons in the input layer. The

output layer of the network contains neurons corresponding to the variables which the network should come up with after completing the calculation. As stated in the model, it is agreed that the output should involve only one variable, therefore the output layer will consist of only one neuron. It is more complicated to determine the amount of neurons in the hidden layer. There are no methods or algorithms for determining their specific number. Determining the optimal (in terms of network utilization) number of neurons in the hidden layer needed for solving the problem of forecasting the levels of PM10 will constitute one of the studies described in the following paragraphs of this work. The model of the structure of the implemented neural network, constructed on the basis of such assumptions, is shown in Figure 7.

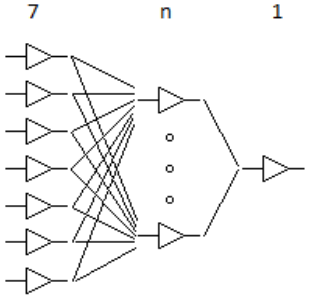

Fig. 7: The applied model of the artificial neural network

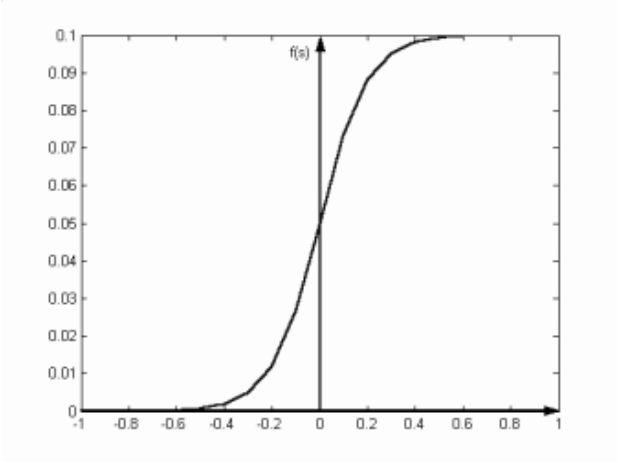

Fig. 8. Unipolar sigmoid activation function

The next stage of constructing the network is the adoption of the activation function. The choice of this function depends on the problem which the network must solve. For multi-layer neural networks, where the output values are expected to be in a specified range, non-linear activation functions are used. Neurons then show the greatest ability to learn. This process involves a smooth mapping of any relationship be-

tween the input and output network variables. There are two main non-linear activation functions [8]: unipolar sigmoid (Figure 8) and bipolar sigmoid (Figure 9).

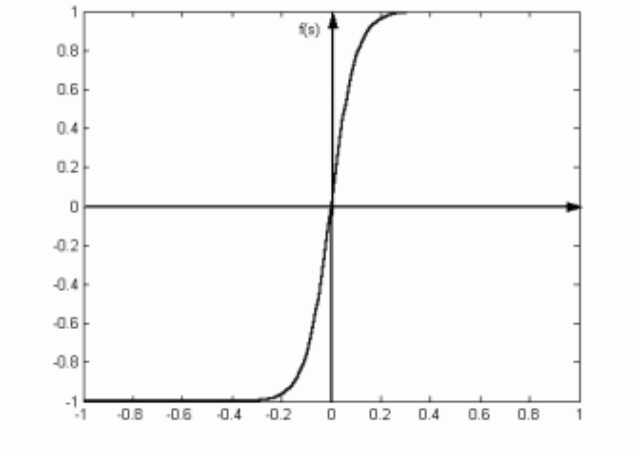

Fig. 9. Bipolar sigmoid activation function

It was decided that, from the point of view of the study, it does not matter whether the unipolar or bipolar sigmoid function is selected. Each of these functions can be easily converted to the other without loss of accuracy. For the constructed network, the bipolar sigmoid function will be used, which takes values in the range of (-1, 1). Based on the above set configuration parameters (unidirectional network, one hidden layer, 7 neurons in the input layer, one neuron in the output layer, bipolar sigmoid activation function), the developed model was implemented with the use of an artificial neural network.

4. THE LEARNING PROCESS OF THE DEVELOPED MODEL

It is assumed that the implementation of the artificial neural network is not sufficient for achieving the correct results (forecasting PM10 levels). For reasoning, we should have structured knowledge [1, 4]. After the model implementation phase using the unidirectional network, it becomes necessary for the network to learn. Doing so involves providing the network with the values of the input and output variables. Entering the values is followed by a process of modifying weights at the connections between neurons in the adjacent layers. The value of the network output variable depends on the values of these weights. This process is repeated until the network calculation error reaches the setpoint (minimum) value.

The supervised learning method was assumed to be carried out for the network learning process. This means that when entering the input values into the network, the

values of the output variables were already predetermined. To obtain those values, the weights were modified within the network structure. To use this method, first the input and output variables which the network will use to learn on must be obtained. Another important requirement, in relation to the values of the variables, was their cyclical one-hour reading on the measuring equipment. This approach resulted from forecasting the level of PM10 in 24-hour cycles. Archival values of PM levels were acquired from the Foundation: Agency of Regional Air Quality Monitoring in the Gdańsk metropolitan area (ARMAAG). The meteorological data was freely available on the Internet. The location of the measurement stations which were the source of the data was also very significant. The station located in the shortest distance from the meteorological measurement station was selected to be the source of data on the particulate matter, which was located on Ostrzycka Street in Gdansk (Szadółki district). The distance between the two stations was 10 km. The data obtained from them covered two full calendar years (2012 and 2013).

After gathering the training data, the best method of learning had to be chosen, which could be applied to the implemented network structure. The basic method of supervised learning for multilayer, unidirectional networks is the BackPropagation method. The BackPropagation algorithm is based on the decline in the value of the sum of squared errors of learning so as to get closer to the assumed minimum value. This process begins by changing the weights on the connections between the output neurons and the penultimate layer of neurons. Then, it heads back toward the input layer. The weights are updated according to the principle of minimizing the error. The speed with which the training error decreases depends on the network learning coefficient. It is often difficult to find the right learning coefficient. The BackPropagation algorithm was used for training the network proposed in this work, which was expanded with the automatic selection of the learning coefficient. Faster learning algorithms were also considered: resilientpropagation and manhattanpropagation. In these algorithms, the weights of the input neurons are not determined by the size of the gradient and the learning coefficient (as is the case in the regular BackPropagation). To change the weights, both methods use the relationship given in formula 1[5]

$$\Delta w_{ij} = -sign(\frac{\partial E}{\partial w_{ij}})\Delta_{ij} \tag{1}$$

When changing the weights, the two methods differ only in terms of the Δij value. To update the manhattanpropagation algorithm, a constant value of Δij is used. In the resilientpropagation algorithm, the Δij value varies and depends on whether the minimum value of the network error was obtained in the last iteration. The effectiveness of training the network with the use of a fixed Δij method - manhattanpropagation, and a variable Δij - resilientpropagation, will also be verified. The final learning algorithm

taken into account is the Scaled Conjugate Gradient algorithm (SCG). It is used to solve problems in large-scale networks. SCG belongs to the class of conjugate algorithms which show a high superlinear similarity in a large number of nonlinear problems [6]. Through the use of SCG, the time-consuming linear search of the per one iteration of learning is avoided. This approach makes the algorithm faster than other learning algorithms, in the case of some problems, and gives better results. The learning algorithms discussed above (BackPropagation, ResilientPropagation, ManhattanPropagation, ScaledConjugateGradient) have been implemented and adapted to cooperate with the proposed neural network. After learning, the knowledge base(neural network) will be built, which will be used for reasoning [7].

5. PM10 FORECASTING PROCESS WITH THE USE OF THE DEVELOPED MODEL

The first necessary step in the network forecasting was to determine the verification method to check the correctness of the results returned by the network. For this purpose, the RealPM10 function was implemented. Its aim is to check the accuracy of forecasts. It converts the values returned by the network (from -1 to 1) to the corresponding levels of PM10. It collects the absolute differences between the expected (real) value and the value obtained by the network, and then calculates their average (mean error). In addition, the maximum and minimum difference is also checked. Errors are reported in the PM10 measurement units - $\mu g/m^3$.

In order to properly check the functioning of the network, it was necessary to prepare a testing set. To achieve this, 48 days were selected from the training set comprising two full calendar years (and removed from the training set). The selection principle was as follows: in each month of the year, two days were selected: one at the beginning of the month (the first day) and one in the middle of the month (the fifteenth day). In this way, a spectrum was obtained covering two full calendar years with an interval of approximately two weeks. Moreover, such hours were selected to cover the scope of 24 hours. All the selected dates are shown in Table 2.

Table 1. Data used in the process of testing the accuracy of the network

12-01-01 01:00	12-01-15 13:00	12-02-01 03:00	12-02-15 15:00	12-03-01 05:00	12-03-15 17:00	12-04-01 07:00	12-04-15 19:00
12-05-01 09:00	12-05-15 21:00	12-06-01 11:00	12-06-15 23:00	12-07-01 13:00	12-07-15 01:00	12-08-01 15:00	12-08-15 03:00
12-09-01 17:00	12-09-15 05:00	12-10-01 19:00	12-10-15 07:00	12-11-01 21:00	12-11-15 09:00	12-12-01 23:00	12-12-15 11:00
13-01-01 01:00	13-01-15 13:00	13-02-01 03:00	13-02-15 15:00	13-03-01 05:00	13-03-15 17:00	13-04-01 07:00	13-04-15 19:00
13-05-01 09:00	13-05-15 21:00	13-06-01 11:00	13-06-15 23:00	13-07-01 13:00	13-07-15 01:00	13-08-01 15:00	13-08-15 03:00
13-09-01 17:00	13-09-15 05:00	13-10-01 19:00	13-10-15 07:00	13-11-01 21:00	13-11-15 09:00	13-12-01 23:00	13-12-15 11:00

Testing the proper operation of the network involved having the network generate output for the selected data (Table 2). Then the RealPM10 function was applied with the generated results as the function input and 3 values of the network error were returned: the average, maximum and minimum error.

After determining the verification method for checking the proper operation of the network, the forecasting process was initiated. The first stage of the study was to choose the best learning methods and the number of neurons in the hidden layer of the neural network. An insufficient number of neurons in the hidden layer is associated with the network having difficulty in learning. The lengthy learning process is caused by problems with the optimal adjustment of weights to a small amount of neurons. On the other hand, if the number of neurons is too high, this may cause the network's inability to generalize its calculations and lead to an increase of the actual error. In both cases, the network learning error falls to a certain level. With seven neurons in the input layer, the reasonable minimum number of neurons in the hidden layer, which is the starting point of the test, is 5.

The upper limit at which the study should be terminated was set at 37 neurons. If the results were inconclusive after the first study, the scope would be extended. The test was run twice and the results were averaged. A 60-minute learning process was carried out for each possible combination of learning methods (BackPropagation, ResilientPropagation, ManhattanPropagation, ScaledConjugateGradient), as well as for the number of neurons in the hidden layer (5 to 37). After that, the network performance was verified with the use of RealPM10. Table 3 and Table 4 show the results for the number of neurons (the 'Neur.' column): 'Error' - network learning error, 'PM10' - the average actual error of PM10, 'MaxPM10' - the maximum actual error of PM10, 'MinPM10' - the minimum actual error of PM10, 'SumPM10' = 'PM10' + 'MaxPM10' + 'MinPM10'. On the basis of the 'SumPM10' column, the configurations giving the smallest errors were marked.

Table 2. Results of the ANN (artificial neural networks) trained with the BackPropagation and ResilientPropagation methods.

BackPropagation						**ResilientPropagation**					
Neur.	Error[%]	PM10	MaxPM10	MinPM10	SumPM10	Neur.	Error[%]	PM10	MaxPM10	MinPM10	SumPM10
5	1.47209	8.39246	30.60900	0.13300	39.13446	5	1.46186	8.38830	37.78150	0.08550	46.25530
6	1.46567	8.19068	30.60900	0.13300	38.93268	6	1.44244	8.46648	37.78150	0.08550	46.33348
7	1.45000	8.20002	30.60900	0.08550	38.89452	7	1.43660	8.52513	43.17750	0.02850	51.73113
8	1.41401	8.06906	29.18400	0.28500	37.53806	8	1.40992	8.00058	30.49500	0.23750	38.73308
9	1.45427	8.16485	33.15500	0.28500	41.60485	9	1.39947	7.99445	31.70150	0.14250	39.83845
10	1.42826	8.22733	33.87700	0.06650	42.17083	10	1.38667	7.96687	31.70150	0.14250	39.81087
11	1.39013	8.02354	31.06500	0.72200	39.81054	11	1.38684	7.95744	34.53250	0.06650	42.55644
12	1.43853	8.19316	31.06500	0.65550	39.91366	12	1.36687	8.00959	34.53250	0.06650	42.60859
13	1.44030	8.09064	31.06500	0.65550	39.81114	13	1.36737	8.22014	36.57500	0.06650	44.86164

14	1.36676	8.00355	30.40000	0.19950	38.60305	14	1.35315	8.09321	29.69700	0.25650	38.04671
15	1.41688	7.97970	30.40000	0.19950	38.57920	15	1.35149	8.09994	31.06500	0.25650	39.42144
16	1.42158	8.02156	30.40000	0.12350	38.54506	16	1.34237	8.03667	31.06500	0.00950	39.11117
17	1.37681	7.82266	27.44550	0.12350	35.39166	17	1.32581	8.06689	35.43500	0.35150	43.85339
18	1.42423	7.96971	32.70850	0.12350	40.80171	18	1.32836	8.24758	35.43500	0.12350	43.80608
19	1.40559	7.93870	32.70850	0.00000	40.64720	19	1.31762	8.16419	35.43500	0.12350	43.72269
20	1.35906	7.78723	29.26950	0.08550	37.14223	20	1.32296	7.81652	31.14100	0.02850	38.98602
21	1.40861	8.05046	29.26950	0.08550	37.40546	21	1.30283	7.95922	39.01650	0.02850	47.00422
22	1.42086	8.07071	31.31200	0.08550	39.46821	22	1.30697	8.03416	39.01650	0.02850	47.07916
23	1.35437	7.90875	32.84150	0.34200	41.09225	23	1.28206	9.04578	49.19100	0.01900	58.25578
24	1.41916	7.97535	32.84150	0.08550	40.90235	24	1.30626	8.71209	49.19100	0.01900	57.92209
25	1.41340	7.96133	32.84150	0.08550	40.88833	25	1.27713	8.55627	49.19100	0.01900	57.76627
26	1.34265	7.78406	31.62550	0.11400	39.52356	26	1.25315	7.86956	29.57350	0.21850	37.66156
27	1.42834	7.99385	32.96500	0.11400	41.07285	27	1.31265	8.07253	33.25950	0.09500	41.42703
28	1.40594	8.04175	32.96500	0.11400	41.12075	28	1.27251	8.15206	35.24500	0.09500	43.49206
29	1.33171	8.11973	29.97250	0.02850	38.12073	29	1.26381	7.48244	40.33700	0.01900	47.83844
30	1.40242	8.00504	29.97250	0.02850	38.00604	30	1.28057	7.90746	40.33700	0.01900	48.26346
31	1.40191	8.04030	29.97250	0.02850	38.04130	31	1.27036	8.24791	40.33700	0.01900	48.60391
32	1.31708	8.22324	33.03150	0.18050	41.43524	32	1.22728	7.78169	31.60650	0.03800	39.42619
33	1.41072	8.02988	33.03150	0.18050	41.24188	33	1.26536	7.98069	31.60650	0.03800	39.62519
34	1.41593	8.09334	33.03150	0.12350	41.24834	34	1.27900	7.82134	31.60650	0.03800	39.46584
35	1.31274	8.37979	30.67550	0.04750	39.10279	35	1.23477	8.46846	47.12000	0.03800	55.62646
36	1.40871	8.28093	34.18100	0.04750	42.50943	36	1.25137	8.10093	47.12000	0.02850	55.24943
37	1.40454	8.16076	34.18100	0.03800	42.37976	37	1.26581	8.11610	47.12000	0.02850	55.26

Table 3. Results of the ANN trained with the ManhattanPropagation and ScaledConjugateGradient methods.

ManhattanPropagation						**ScaledConjugateGradient**					
Neur.	Error[%]	PM10	MaxPM10	MinPM10	SumPM10	Neur	Error[%]	PM10	MaxPM10	MinPM10	SumPM10
5	1.46782	8.62897	32.63250	0.15200	41.41347	5	1.46820	8.34060	30.29550	0.38000	39.01610
6	1.46310	8.46767	32.63250	0.15200	41.25217	6	1.47201	8.35129	32.56600	0.27550	41.19279
7	1.45256	8.42307	37.65800	0.15200	46.23307	7	1.45451	8.29977	32.56600	0.27550	41.14127
8	1.42992	8.19533	34.80800	0.09500	43.09833	8	1.41893	8.08015	29.62100	0.36100	38.06215
9	1.42301	7.92547	34.80800	0.06650	42.79997	9	1.46651	8.20186	32.57550	0.36100	41.13836
10	1.40405	7.90301	34.80800	0.06650	42.77751	10	1.45193	8.28941	32.57550	0.00950	40.87441
11	1.40316	7.85294	30.13400	0.68400	38.67094	11	1.42090	8.05640	33.07900	0.06650	41.20190
12	1.41630	8.17772	32.91750	0.68400	41.77922	12	1.45257	8.19494	33.07900	0.06650	41.34044
13	1.39579	8.08034	32.91750	0.20900	41.20684	13	1.45991	8.20068	33.07900	0.06650	41.34618
14	1.37476	7.66353	29.26000	0.12350	37.04703	14	1.46699	8.38692	31.83450	0.52250	40.74392
15	1.38287	7.78426	30.61850	0.12350	38.52626	15	1.47146	8.43254	34.55150	0.13300	43.11704

16	1.39847	7.96562	30.61850	0.12350	38.70762	16	1.46123	8.29746	34.55150	0.13300	42.98196
17	1.34958	8.32180	28.85150	0.86450	38.03780	17	1.43451	7.94913	30.43800	0.07600	38.46313
18	1.39293	8.18316	36.14750	0.04750	44.37816	18	1.60173	8.38543	30.43800	0.07600	38.89943
19	1.39191	8.17501	36.14750	0.04750	44.37001	19	1.48138	8.45210	34.20000	0.07600	42.72810
20	1.34560	7.98297	33.27850	0.09500	41.35647	20	1.41985	8.23017	30.78950	0.85500	39.87467
21	1.38652	7.98366	33.27850	0.09500	41.35716	21	1.47604	8.32606	30.78950	0.32300	39.43856
22	1.40329	8.07962	33.27850	0.09500	41.45312	22	1.48738	8.35789	30.78950	0.21850	39.36589
23	1.32823	7.94259	30.54250	0.05700	38.54209	23	1.45109	8.33981	34.50400	0.49400	43.33781
24	1.39608	8.03433	30.54250	0.05700	38.63383	24	1.50453	8.35159	34.81750	0.29450	43.46359
25	1.37739	8.05422	30.54250	0.05700	38.65372	25	1.49513	8.36858	34.81750	0.29450	43.48058
26	1.30343	7.85729	31.07450	0.05700	38.98879	26	1.43127	8.33566	30.93200	0.37050	39.63816
27	1.38882	7.83710	32.69900	0.05700	40.59310	27	1.50023	8.33734	31.52100	0.22800	40.08634
28	1.38934	7.84331	32.69900	0.02850	40.57081	28	1.48660	8.32965	31.52100	0.22800	40.07865
29	1.32295	7.71143	30.02000	0.02850	37.75993	29	1.42375	8.24818	32.25250	0.16150	40.66218
30	1.38002	7.80692	31.52100	0.02850	39.35642	30	1.51721	8.42264	32.25250	0.16150	40.83664
31	1.37025	7.86521	33.78200	0.02850	41.67571	31	1.50607	8.43283	35.91000	0.16150	44.50433
32	1.29550	8.15674	33.17400	0.04750	41.37824	32	1.44027	8.39622	31.30250	0.24700	39.94572
33	1.39937	8.16584	33.17400	0.04750	41.38734	33	1.52094	8.40562	31.30250	0.24700	39.95512
34	1.39496	8.09605	33.17400	0.04750	41.31755	34	1.51725	8.40176	38.41800	0.24700	47.06676
35	1.31573	8.23393	32.57550	0.00950	40.81893	35	1.45979	8.70932	32.84150	0.14250	41.69332
36	1.39603	8.14457	32.57550	0.00950	40.72957	36	1.55040	8.70724	35.00750	0.08550	43.80024
37	1.38762	8.04373	32.57550	0.00950	40.62873	37	1.53416	8.42571	35.00750	0.08550	43.51871

Comparing the results from Table 3 and Table 4, it can be concluded that the BackPropagation learning method returns the smallest average error among all the tested methods. The average errors are as follows: BackPropagation - 39.7297, ResilientPropagation - 45.4439, ManhattanPropagation - 40.7727, ScaledConjugateGradient - 41.3029. It was decided that the BackPropagation method will be used for further research.

The next stage of research was to determine the number of neurons in the hidden layer. Analyzing the results obtained by the network trained with the BackPropagation method, there were errors generated by the network with 8, 17, 20, and 21 neurons in the hidden layer. They are lower than the errors for other structures (other numbers of neurons in the hidden layer). In particular, a network with 17 neurons is worth noting - total error = 35.39166, much lower than any other. We can therefore conclude that an effective network structure is trained with the BackPropagation method with 17 neurons in the hidden layer (average PM10 error = 7.82266, maximum PM10 error = 27.44550, minimum PM10 error = 0.12350).

After selecting the number of neurons in the hidden layer, the suggested network had yet to be tested in terms of forecast error. Table 5 shows detailed stages of testing the network.

Table 4. Testing the network after the BackPropagation learning method.

Date	Fore.pm10	Real.pm10	Difference	Date	Fore.pm10	Real.pm10	Difference
12-01-01 01:00	11.286	4.503	6.783	13-01-01 01:00	12.9675	18.297	5.3295
12-01-15 13:00	13.908	19.703	5.795	13-01-15 13:00	25.783	31.901	6.118
12-02-01 03:00	33.934	6.6025	27.3315	13-02-01 03:00	11.97	9.101	2.869
Explanation: drop in wind + night (light traffic)				13-02-15 15:00	21.489	17.898	3.591
12-02-15 15:00	37.2685	33.896	3.3725	13-03-01 05:00	16.1785	11.9035	4.275
12-03-01 05:00	11.3905	5.301	6.0895	13-03-15 17:00	14.364	1.7955	12.5685
12-03-15 17:00	12.977	12.502	0.475	13-04-01 07:00	14.269	22.097	7.828
12-04-01 07:00	10.564	6.6975	3.8665	13-04-15 19:00	22.648	15.7985	6.8495
12-04-15 19:00	12.1125	14.801	2.6885	13-05-01 09:00	28.234	47.5	19.266
12-05-01 09:00	14.269	21.204	6.935	13-05-15 21:00	14.1455	14.2975	0.152
12-05-15 21:00	21.717	3.496	18.221	13-06-01 11:00	12.5115	21.8975	9.386
Explanation: increase in temp by 5deg. in 4h+drop in wind				13-06-15 23:00	22.0305	12.7015	9.329
12-06-01 11:00	11.533	15.903	4.37	13-07-01 13:00	15.903	13.604	2.299
12-06-15 23:00	15.3425	15.0005	0.342	13-07-15 01:00	11.2195	5.301	5.9185
12-07-01 13:00	23.579	15.2	8.379	13-08-01 15:00	15.6655	9.6995	5.966
12-07-15 01:00	8.4645	2.603	5.8615	13-08-15 03:00	10.944	10.0035	0.9405
12-08-01 15:00	18.772	21.4035	2.6315	13-09-01 17:00	10.5735	3.002	7.5715
12-08-15 03:00	14.9625	10.298	4.6645	13-09-15 05:00	13.3285	3.6955	9.633
12-09-01 17:00	12.7015	6.403	6.2985	13-10-01 19:00	19.3135	6.6975	12.616
12-09-15 05:00	9.177	9.3005	0.1235	13-10-15 07:00	16.3875	32.1005	15.713
12-10-01 19:00	21.6505	49.096	27.4455	Explanation: no wind + morning (heavy traffic)			
Explanation: increased wind + direction from city				13-11-01 21:00	14.9055	3.496	11.4095
12-10-15 07:00	15.6845	13.3	2.3845	13-11-15 09:00	18.1925	15.0005	3.192
12-11-01 21:00	13.509	23.3035	9.7945	13-12-01 23:00	8.455	13.8035	5.3485
12-11-15 09:00	24.9185	49.5995	24.681	13-12-15 11:00	16.188	5.1015	11.0865
Explanation: clouds disappeared + morning (heavy traffic)						**Average error**	7.82265
12-12-01 23:00	19.969	23.199	3.23			**Average error without indicated dates**	5.89588

On the basis of the results shown in Table 5, the cases when the value of the PM10 forecast error was high were selected and analyzed. Some interesting situations when the increased value of the forecast error occurred include these days and times: 12-02-01 03:00, 12-05-15 21:00, 12-10-01 19:00, 12-11-15 9:00, 13-10-15 7:00. A meteorological analysis of these cases followed. After examination, it turned out that the causes refer to phenomena which are very difficult to predict. In these cases they involved for instance an increase in wind power, or snow occurring quite suddenly. On the basis of measurements taken over 24 hours before, they were difficult to predict. After eliminating these cases, the average forecast error decreased from 7.82265 to 5.89588 and the maximum value of the forecast error fell from 27.44550 to 12.5685.

6. SUMMARY

The purpose of the work was to build a model to predict the level of PM10 in urban areas. In the model construction process, first the usefulness of the input variables for the model was assessed, and then a unidirectional neural network was used for implementing the model. The choice of a unidirectional neural network was due to its possibility of solving nonlinear prediction problems with plenty of independent input variables.

The selection of the network was followed by the construction of its structure. It was decided that a three-layer network with one layer of hidden neurons will be used. The choice of only one layer of hidden neurons was due to the nature of the problem to be solved, namely the forecasting of the PM10 level. There is no need to use a larger number of hidden layers for this type of problem. The chosen activation function was bipolar sigmoid, because then neurons show the greatest ability to learn. This function also allows for smooth mapping of any relationship between the input and output variables.

The next step was to develop a method for testing the network. To do so, the results for dates evenly spread over two years were calculated, and the results were given in real PM10 values.

After selecting the network structure and the testing method, the learning algorithm was chosen. The following algorithms were taken into account: BackPropagation, ResilientPropagation, ManhattanPropagation, ScaledConjugateGradient. BackPropagation was eventually applied for the learning process due to the smallest learning error returned by the network.

After selecting the network learning method, the optimal number of neurons in the hidden layer was determined. 5 to 37 neurons were considered. A neural network with 17 neurons in the hidden layer was selected, as it showed the smallest forecast error.

In the next stage, the results of testing were examined in detail. It was concluded that the higher forecast errors are caused by sudden changes in weather conditions that were difficult to predict in the past.

Using the developed model, an average forecast error equal to 7.82265 $\mu g/m^3$ could be obtained.

The most labor-consuming stage in the construction of the network was the stage of learning and collecting results, because learning takes time, and the training configurations were numerous. The selection of input variables proved to be a relatively simple stage, as most of the variables were analyzed in a similar manner and the process ran smoothly.

With the use of a neural network, the future value of PM10 can be predicted, although it is worth considering how to improve the process. There are plans to improve the functioning of the network through adding neurons in the input layer. These neu-

rons would be responsible for forecasting the meteorological conditions, derived from specialized models for predicting the weather. The following process in building a model, in order to increase the accuracy of calculations, would be to increase the size of the training set.

7. REFERENCES

[1] Czarnecki A., Orłowski C., Sitek T., Ziółkowski A.*: Information Technology Assessment Using a Functional Prototype of the Agent Based System*// Foundations of Control and Management Sciences. - 2008, no 9, s. 7–28 : 5 rys., 5 tab. - Bibliogr. 8 poz. - ISSN 1731-2000

[2] Drzeniecka-Osiadacz A., Netzel P.: *Wpływ warunków meteorologicznych oraz cyrkulacji atmocferycznej na stężenie PM10 we Wrocławiu*, 2010.

[3] Kosiński R.,: *Sztuczne sieci neuronowe Dynamika nieliniowa i chaos*, 2005.

[4] Kowalczuk Z., Orłowski C.: *Design of Knowledge-based Systems in Environmental Engineering, Cybernetics and Systems*, vol. 35 nr 5–6, 2004, s. 487–498. 19 rys. 1 tab. bibliogr. 11 poz.

[5] Neutelings J.,: *SmarTee: A neural network based navigational bot*, 2011.

[6] Peterson C., Rögnvaldsson T., Lönnblad L.,: *JETNET 3.0: A Versatile artificial neural network package*, 1994.

[7] Sitek T., Orłowski C.: *Model of Management of Knowledge Bases in the Information Technology Evaluation Environment* // W: Information systems architecture and technology : models of the organisation's risk management / editors: Zofia Wilimowska [et al.]. - Wrocław 2008 : Oficyna Wydawnicza Politechniki Wrocławskiej, 2008. - (Biblioteka Informatyki Szkół Wyższych). - S. 221–231 : 5 rys. - Bibliogr. 9 poz. - ISBN 978-83-7493-413-8

[8] Tadeusiewicz R., *Wprowadzenie do praktyki stosowania sieci neuronowych*.

Knowledge Management, Decisional DNA,
Set of Experience Knowledge Structure, R&D Organizations

Norbert BOLEŃSKI*, Edward SZCZERBICKI**

KNOWLEDGE-BASED MANAGEMENT AS THE KEY SUCCESS FACTOR FOR RESEARCH AND DEVELOPMENT ORGANIZATIONS

A key to achieving success in project organizations lies in exemplary management of processes within those organizations, while the ongoing projects are mainly characterized by their uniqueness. The situation is no different in commercial research and development organizations (R&D) where accuracy and repeatability of elementary processes guarantees efficient and productive realization of all enterprises. R&D organizations are entities whose main objective is to carry out basic research as well as experimental development works, and to distribute the results via didactic activity, publications or transfer of technology. The aim of the present paper is to outline the area of knowledge management within this type of organizations and to describe the model of management.

1.NEWLY ARISING NEEDS OF TODAYS CUSTOMER

Over the years, along with changes in the concept of management, the configuration of organizational units within the company have been constantly evolving, resulting in the creation of new organisational structures. Previous linear and functional structures are being replaced by matrix, divisional, and nowadays more often – project focused structures. Therefore, these latter deserve a special attention in today's world where the relation between the customer and the company is constantly changing [1].

Project organizations are organizations in which each task is carried out individually. They are characterized by a lack of repeatability due to the necessity of matching specific solutions to specific customer.

* Gdansk University of Technology, Gdansk, Poland.
** The University of Newcastle, Newcastle, Australia

Research and development organizations (R&D), which are a particular form of project organization, require special attention. In this case, likewise, all tasks can be treated as separate projects. However, when it comes to R&D organization, no less important than the same execution of a task is to mark company's contribution to the development of the sector or industry. For this purpose, R&D organization share part or all of their results and transfer technology and knowledge of the area of the research. What distinguishes the R&D organization from other project organizations is a uniqueness of results, even when repeating the same action. The complexity of this problem, have been mentioned by authors in previous publications [2].

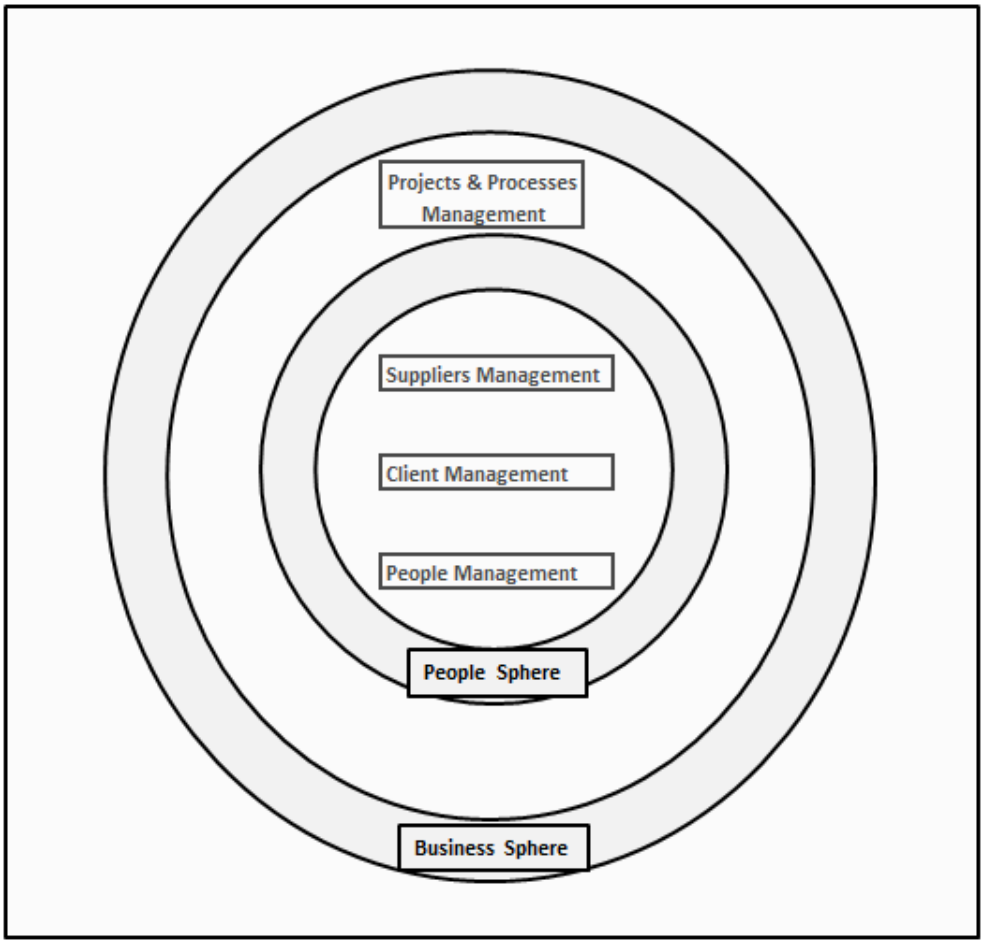

Fig. 1. Interpenetration of company business spheres

Nowadays the customer is "the boss" of any company. Various processes associated with globalization cause the problem of shortage of customers. Rather than a shortage of goods, which only a few decades ago was a major issue that companies had to face, in todays' market, the challenge is to attract the customer and convince them to choose one's offer. This is caused by the fact that we live in an era of almost unlimited choice of products and services. No longer it is possible to satisfy the customer offering only a product. More and more often we can observe increasingly common situation when a company is losing satisfied customers to competitors, which appear to be even better / cheaper / more convenient. In other words, to attract a customer an organization needs to be able to offer him additional value. Each client requires an individual approach, and the products or services offered must, in all details, respond to his needs. Customers desire products and services that simplify and improve their lives [3]. According to the definition, R&D organizations due to their characteristics, especially contributing to the development of the business, are the best example of a modern company. Unfortunately, referring to Phillip Kotler's words, most companies tend to focus their actions on acquiring new customers and increasing their profits [4]. This trend is even more objectionable when it comes to companies operating within R&D business. Taking into account the nature of the industry and what results from its' specifics, limited group of target customers, it is highly unlikely an approach that leads to a long lasting success.

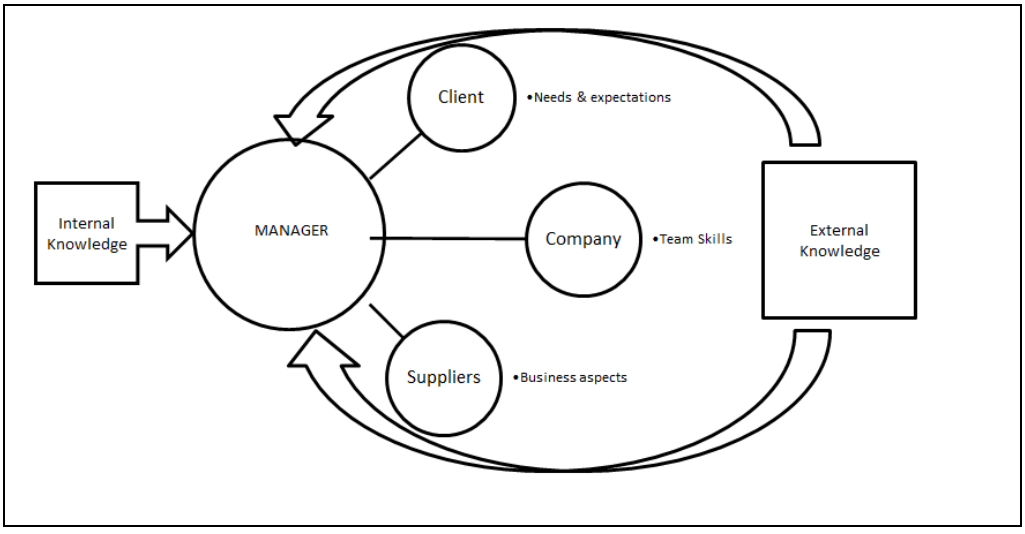

Fig. 2. The flow of knowledge in the organization environment

Evidently, contemporary customer requires a full attention. This determines the need to improve quality of products, so as not to be left out from the competition. In consequence, companies are forced to work on an improvement of management processes within its' structures. The process focused approach enables a holistic improvement of the elements that add value cross-sectional across the entire organization. Obviously, process based management is not a novelty, and is closely related to the quality management. Already at the beginning of the twentieth century, the Ford Motor Company introduced the control of production processes. This was to verify the quality of products that had been sold by the company. Products that met the quality criteria were sold, products of lower quality were revised or offered to the customers at a discounted price. Testing the final product was the first stage of quality management. The second was controlling the quality by probing not only the final product but also all manufacturing procedures, by verifying the knowledge and skills of the employees. The third stage was offering customers products that meet established standards, so that they know what to expect in the future. What is more, books of quality had been published. They included required parameters and technological minimums that the product must met in order to satisfy the customer. The final, fourth stage of quality management development is a comprehensive quality management known as TQM (Total Quality Management). The major assumption of this approach is a continuous improvement of processes at each level of the company's activity including the stage of cooperation with both, suppliers and customers. Speaking of TQM, it seems essential to refer to the 14 principles of Deming described in his work "Out of Crisis" published 1986. In the publication mentioned above, the great emphasis is being put on processes and their improvement with simultaneous resignation from the multitude of control. The process is a chain of value creation for both, the customer and the company. Deming also suggests eliminating all barriers within the company, negates separating employees due to the level and division. According to his work, the whole company should be one, compact and constantly improving organism [5]. In today's world, where access to knowledge and technology is almost unlimited and, through TQM processes companies are able to produce products of comparable quality, organizations need to focus on the customer and his satisfaction, not only on the product. Thus, it can be concluded that it changed the definition of success. Success can no longer be a target built only on providing products compatible with the expected specification. Survival and growth in the highly competitive market can be achieved by maximizing the benefit of all stakeholders, and even the environment. This results in the interpenetration of outside business spheres such as the social and cultural (the sphere of human relations) with a strictly business sphere.

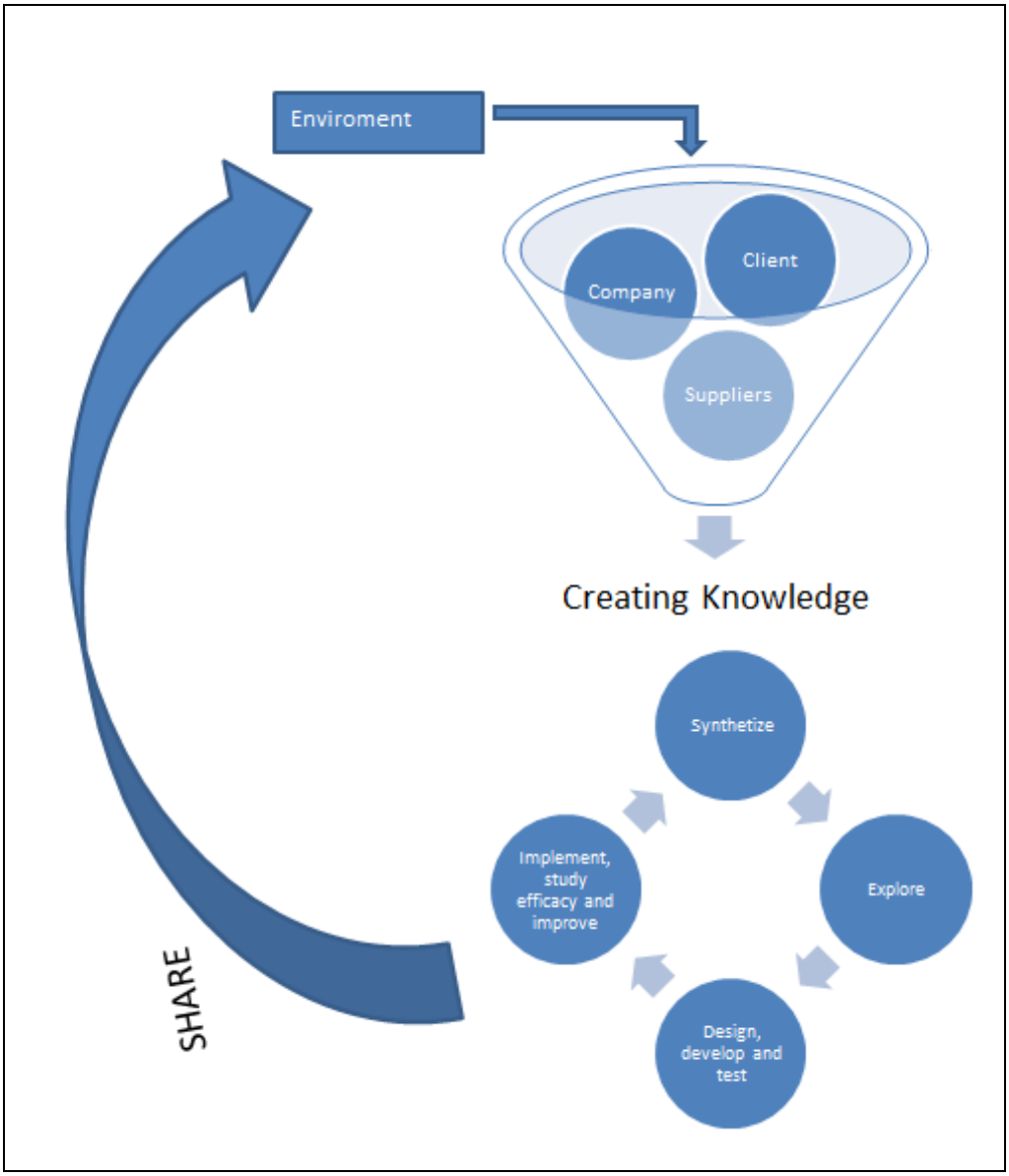

Fig. 3. Process of gathering and creating knowledge

It is an undisputable fact, that organizations such as the R&D require highly skilled personnel. However, due to the formulation of procedures tailored to the main business, it is possible to focus their work around the development of new

pathways that lead to accomplishing the task, and also to partial supervision of the secondary works. Creating new paths requires close cooperation with the customer, which, in turn, requires certain interpersonal skills that allow to fully understand their needs. In other words, there has been an evolution at the level of the desired employee profile - from an expert in the given field for an efficient project manager, acting as a link between the customer, the company and suppliers.

2.A WAY TO ACHIEVE SUCCESS

One of the key factors determining the success of R&D organization is an added value that can be provided to the environment. The literature support readers with a large number of works that analyse key success factors recognised in the projects from different sectors. The most frequent, and at the same time related to the characteristics of R&D organizations are:

- Properly planned resources
- The human factor – appropriate managing of employees, setting clear objectives and indicating the direction of their implementation
- Recognition of the required technology and the selection of tools for the project
- Development of a standard methodology supported by testing and research
- Preparation of financial analysis and accurate estimation of the risk associated to the failure of the project
- Completing the right team to work both primary (experts) and secondary (less experienced staff)
- Selection of a leader

Particular attention should be paid to the point concerning technology and tools, including knowledge management tools. The creation of such a model holds great significance, as the proposed solution could be implemented in other organizations. Currently the most common knowledge management models in enterprises are those based on the concepts of Decisional DNA (DDNA) and the semantic web.

DDNA is a concept in which the inference is based on the result of pre-recorded and performed procedures. Knowledge and experience is used as a result of the analysis carried out by the software. The basis of the concept of DDNA a model of relations between rules, variables, constraints and functions. The combination of these 4 elements allows to save the Set of experience (SOE). SOE

Group within a single structure (f.ex. department in the company) is a Decisional Chromosomes. Groups of chromosomes are called Decisional DNA (DDNA) [6].

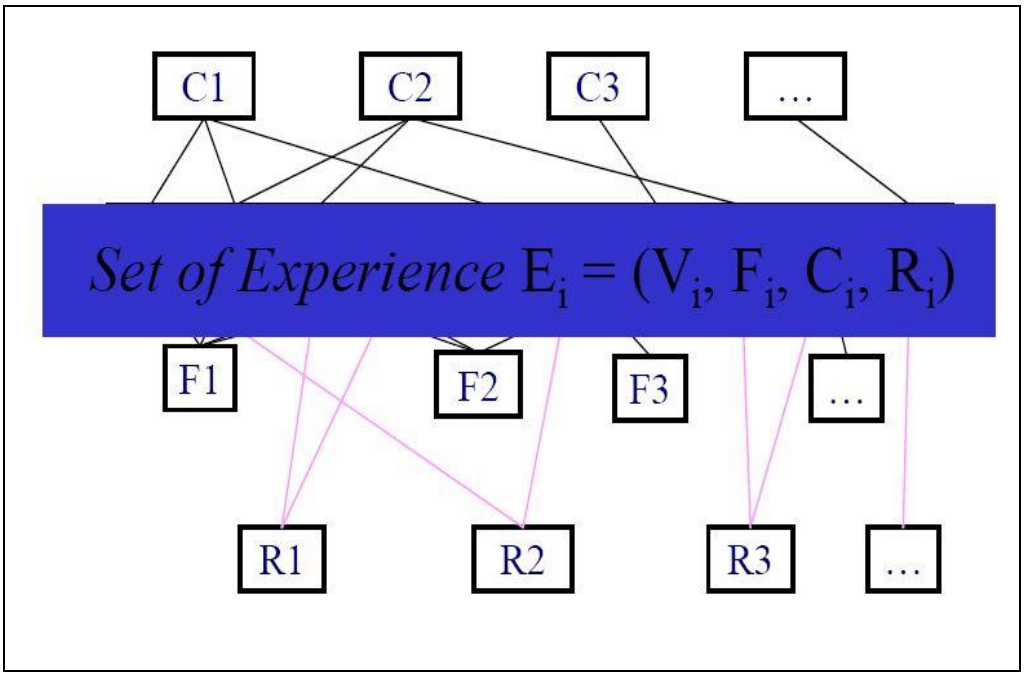

Fig. 4. Set of Experience Knowledge Structure (SOEKS) [6].

Another tool that can be used in knowledge management is a semantic network. In 1999, Atim Berners-Lee wrote: „I have a dream for the Web [in which computers] become capable of analyzing all the data on the Web – the content, links, and transactions between people and computers. A "Semantic Web", which makes this possible, has yet to emerge, but when it does, the day-to-day mechanisms of trade, bureaucracy and our daily lives will be handled by machines talking to machines. The "intelligent agents" people have touted for ages will finally materialize" [7]. Due to the versatile possibilities which are inherent to XML, it was possible to describe the ontology, which on further stage of modeling enabled the recording of knowledge and conclusions.

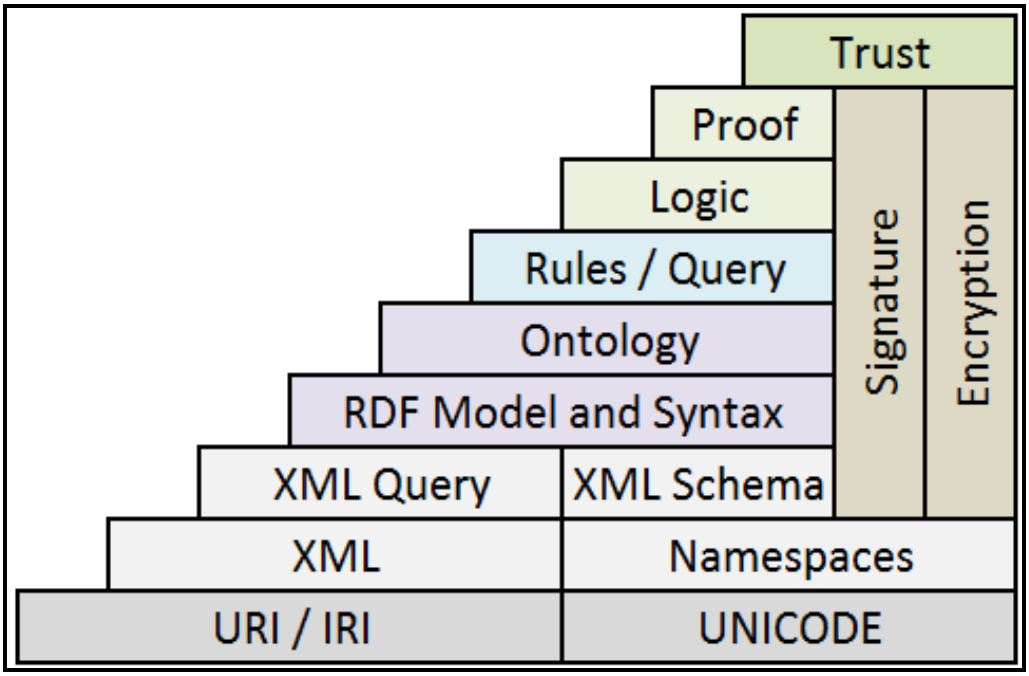

Fig. 5. Semantics technologies

3.CONCLUSION

Summarizing the above considerations on the model of knowledge management in R&D organization issues of benefits and problems associated to the implementation of such a model should be addressed. Regardless of the method of modelling of knowledge that an organization adopts, the aspects of implementation remain similar and will require additional study. This, in turn implies the possibility of resistance of employees. Each person in the organization must be convinced that the model that allows the collection of knowledge is necessary for work. A tool that provide development, both personal and organizational. Growth of knowledge measures based, for example on learning or experience curves [8], as well as HR programs stimulating employees for personal development, inter alia by providing a suitable environment appears to be extremely effective to achieve it.

The implementation of an intelligent model of knowledge management can also bring many purely material benefits. This will allow to transfer some duties to the lower level employees and consequently will significantly reduce costs and allow specialists to focus on more complicated tasks. Models are, by definition, learning models, which results in shortening of the conceptual work on next project. And above all, such model can provide what is core of R & D activities, the development of science, by publishing part of that knowledge to a broad audience. It is the author's opinion, these are some important arguments in favor of implementing an intelligent model of knowledge management in organizations that should convince their senior management to the idea, even though implementation will be associated with additional costs.

It is difficult not to mention the five disciplines proposed by Peter M. Senge [9], which, as the necessary elements, most accurately define a model learning organization, and this is what should the R & D organization be. The important role of the project manager should be emphasized once again. Project manager organize work, selects project group members and should be able to convince the people associated with the project to work on all disciplines, while respecting the soft management issues such as philosophy and culture of the organization, as well as finding the right balance between focusing on learning organization and a model of knowledge (training, etc.), and receiving signals from the environment (implementation of design intent from the client).

REFERENCES

[1] *Struktury organizacyjne*, http://www.zarzadzanie-organizacja.tosi.pl/?p=p_11&sName=struktury-organizacyjne. [acces: 2014-03-01].

[2] BOLEŃSKI N., SZCZERBICKI E., *A concept of using knowledge models to secure intellectual capital in enterprise*, in *Intelligent Information Systems, Knowledge Discovery, Big Data and High Performance Computing.* Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2013, pp. 13–22.

[3] KOTLER P., *Marketing 3.0*. MT Biznes 2010, Warszawa 2010, pp. 17–21.

[4] KOTLER P., *Marketing od A do Z*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2004, p 60.

[5] DAHLGAARD J.J., KRISTENSEN K., KANJI G.K. *Podstawy zarządzania jakości*, Wydawnictwo Naukowe PWN 2004, pp. 17–21.

[6] SANIN C., SZCZERBICKI E., *Application of a Multi-domain Knowledge Structure: The Decisional DNA, in Intelligent Systems for Knowledge Management*, N. Nguyen and E. Szczerbicki, Editors. 2009, Springer Berlin / Heidelberg. pp. 65–86.

[7] BERNERS-LEE T., FISCHETTI M., *Weaving the Web*, Harper, San Francisco 1999. Chapter 12.
[8] PENC J., *Zarządzanie oparte na wiedzy*, in *Organizacja i zarzadzanie*, nr 1(107), 2002, pp. 3–16.
[9] SENGE P.M., *The fifth discipline*

Ontology, optimization.

Maria GALANT-PATER, Anna LAMEK,
Maciej MALARA, Bogusz PRZYBYSŁAWSKI,
Witold REKUĆ, Leopold SZCZUROWSKI*

ONTOLOGY AS A TOOL FOR MODELLING BUSINESS PROCESSES OF TRANSPORT

The subject of this study is an analysis of the solutions in the scope of transportation ontology available in literature. The need for such an analysis results from a broader task being solved by the authors, namely, the development of a methodology for identification and analysis of problems associated with the provision of transport services, as well as a methodology for finding solution to such problems. The study reviews the solutions, concerning the ontology of road and rail transport services known from literature, and evaluates them from the viewpoint of their usability in preparing such products of the methodology being developed as: business motivation model for transport services, business area model, model of organizational structures in transport management, model of processes and definitions of transport management problems. From the viewpoint of a broader context of this study, it is important to analyse transport in many aspects in order to formulate the widest possible range of problems occurring in the provision of transport services. The review of literature carried out as a part of this study shows that with respect to transportation ontology the offer is quite wide and largely covers the area of interest of the authors.

1. INTRODUCTION

This work is directly associated with the task that consists in developing a methodology for identifying and analyzing the problems of transport services provision, and finding solutions to these problems by using a specific IT platform. The development of this methodology requires finding, among other things, ways to model the processes of transport services provision and ways to use them for finding particular solutions to these problems. Modern approaches to solving problems associated with provision of transport services should be based on the use of widely understood knowledge of the subject domain. Therefore, a special attention in our approach is paid to ontology [Gruber, 1995].

* Wrocław University of Technology, Faculty of Computer Science and Management, Wyb. Wyspiańskiego 27, 50-370 Wrocław

The development of the above referred methodology requires multi-aspectual divisions and classifications of concepts associated with this type of activity, as well as depiction of their significant relationships. These concern, in particular: business motivation for provision of transport services, especially a business model of such services; organizational structures in provision of transport services (traffic and its management); transport management processes and problems;impact of business motivation on organizational and process solutions for provision of transport services, including the impact on problem solving.

Bearing in mind the main purpose of our work, it seems that it is important to review the published solutions in the scope of transportation ontology, which is the subject of this paper. This review should indicate the existing proposals, and areas that require additional studies.

2. ONTOLOGIES OF THE TRANSPORT DOMAIN

The review of literature of several past years showed that constructing ontologies has been very popular. Performed were studies on rather general ontologies, but also a lot of effort was put in the development of domain or task ontologies. The work [Fox, Gruninger, 1998] includes so-called Toronto Virtual Enterprise ontologies. The authors of this study determined a network of interrelated ontologies in an enterprise. The transportation ontology associated with the resource ontology is also included in these ontologies. Ontologies addressed strictly to transport can be found in several other items of literature, discussed below.

The work [InteGrail,2010] outlines the Railway Domain Ontology (RDO). Unfortunately, this material does not contain detailed specifications. Whereas, the work [Becker, Smith, 1997] presents the ontology of multi-modal transport system with a quite good level of detail. This study was carried out in the Institute of Robotics at the Carnegie Mellon University in Pittsburgh, and was addressed to the US Army. The subject of the ontology is a US military transport system, in which multi-modal operations, terminal operations and traffic control operations are carried out.

For figure Multi-modal operations shall be understood as air, land and water transport operations. They are accompanied by terminal operations, i.e. operations at the points of loading/unloading. Operations for controlling the traffic of means of transport and controlling terminals are also introduced. These are:

a) planning,
b) apportioning transport,
c) allocating transport,
d) deconflicting priorities,
e) validating transport requests,
f) coordination,

g) in-transit visibility,
h) force tracking.

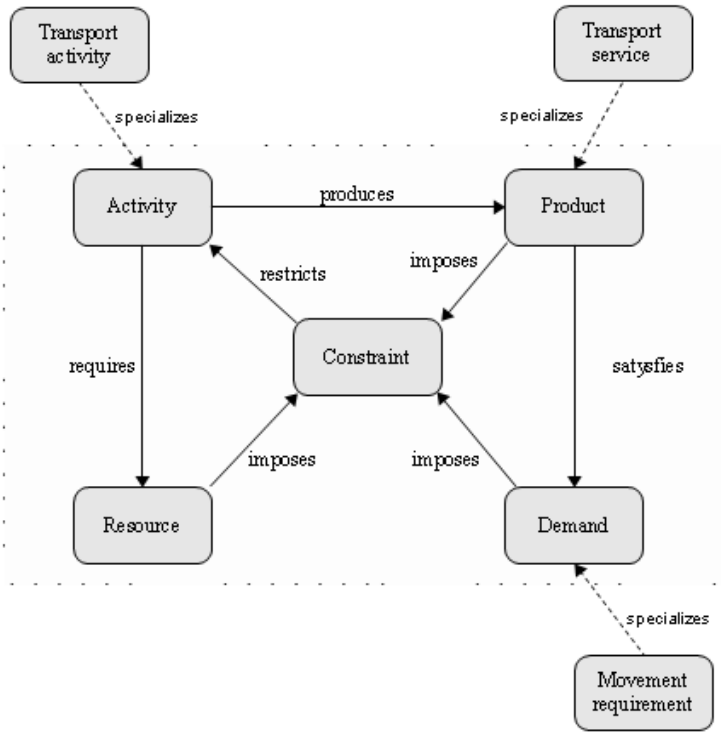

Fig. 1. An outline of general concepts of the multi-modal transport ontology
Source: Own study based on [Becker, Smith, 1997]

For the purposes of building an ontology for a transport system, the authors use a general abstract model of the domain presenting five concepts (see Fig. 1): activity, product, constraint, resource and demand, and then specialize them by introducing the concepts of transport activity, transport service and movement requirement. Movement requirement results from the needs of the time of peace and the time of war. The study mentioned above includes definitions of many other concepts, which are specializations of the concepts from Fig. 1, and therefore deserves special attention. Although it is addressed to a military institution, it may be, in a certain degree, adapted to civilian applications.

Another interesting item is the study [Bauereiss et al. 2012] carried out as part of the "European e-Freight capabilities for Co-modal transport (e-Freight)" project[1].

[1] http://www.efreightproject.eu/ (27.08.2013)

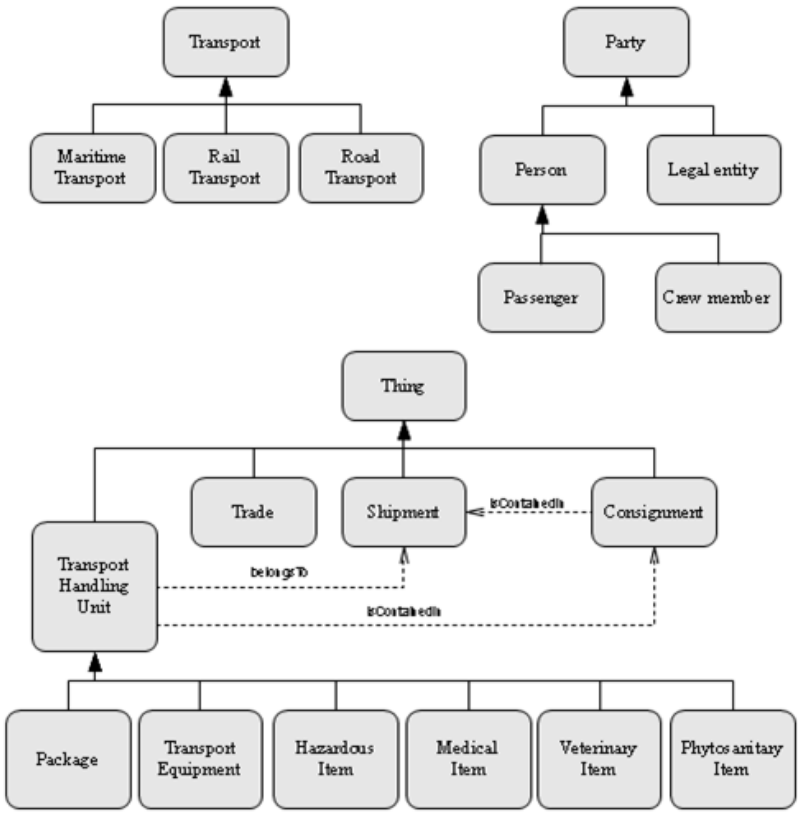

Fig. 2. The main concepts of transportation ontology in the e-Freight project
Source: Own study based on [Bauereiss et al., 2012]

This work concerns the transport of goods and it shall be possible, to a large extent, to use the ontologies presented therein for building the ontology within our task. Concepts of ontologies have been classified in the following groups: Document, Location, Measure, Party, Trade, Transport, TransportEvent, Transport Handling Unit and Other. The main concepts of ontology are shown in Fig. 2

In the ontology presented, different transport events were described with a quite good level of detail. This is shown in Figure 3.

In addition to the concepts, the study presents definitions of properties, as well as information about associations between the concepts. The work is specified in the RDF language.

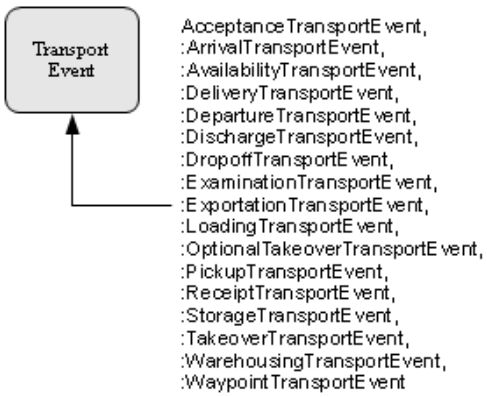

Fig. 3. Classification of transport events in the e-Freight project
Source: Own study based on [Bauereiss et al., 2012]

3. TASK ONTOLOGIES

The studies [Marçal de Oliveira et al, 2013] and [Houd et al, 2010] present transportation ontology from the viewpoint of travel planning and personalization of the interface of a travel planning application. A diagram of the transport ontology is shown in Fig. 4

The work [Sadeghi-Niaraki, Kim, 2009] concerns route planning oriented at individual travellers. The importance of the concept of road segment impedance is emphasized. The road segment impedance is a measure of the cost or expected resistance in connection with the travel from junction A to junction B in a road network. The higher the impedance, the greater the resistance to traffic [Sadeghi-Niaraki, Kim, 2009]. The authors present an ontology associated with the road segment. Basic concepts in this ontology include, in addition to the concept of "road segment", also the concepts of criteria for its evaluation, grouped as:

- qualitative: road, tourist, user, security, equipment and weather criteria,
- quantitative: road, user and vehicle criteria.

Further, the authors present a proposal of a road segment ontology, which includes, e.g., concepts of the road segment, road network, vehicle, and driver, as well as concepts associated with the above-mentioned criteria. The whole is oriented towards building a decision support system for selecting the route for a user travelling in a certain road network, taking into account a wide range of criteria.

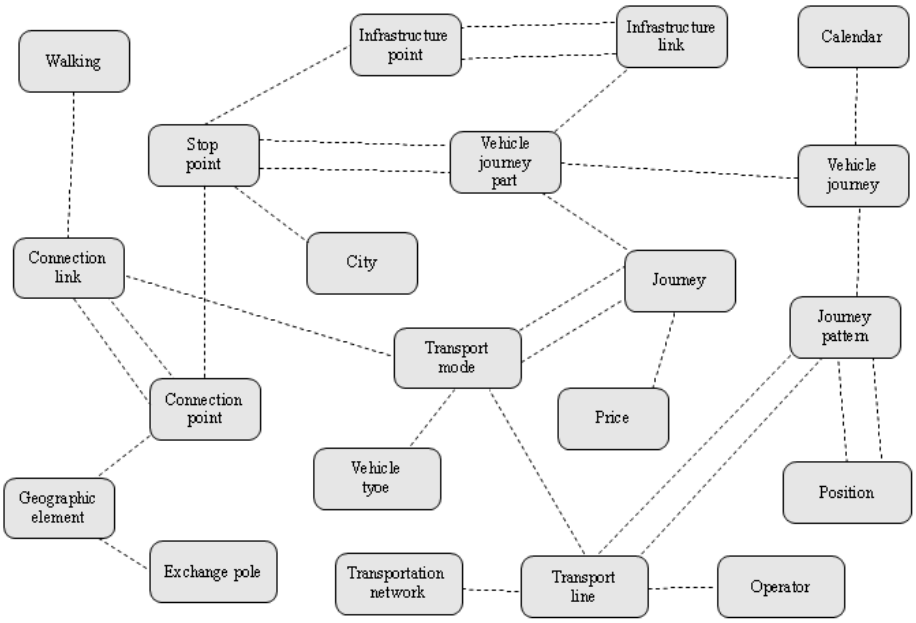

Fig. 4. Diagram of transportation ontology
Source: own study based on [Marçal de Oliveira et al, 2013; Houd et al, 2010]

The work [Samper et al, 2006] discusses an ontology built for the purpose of developing information systems for travellers. The basic subject of the ontological modelling is road traffic. The work does not contain any detailed information on the ontology. However, it includes a general division of the subject of modelling into so-called sub-domains: roads and their classification, vehicles and their classification, location, geography, events, people, routes.

The study [Wang, Ding, Jiang, 2005] proposes an ontology for public urban transport. The aim of development of this ontology is to solve the problem of providing passengers with the access to an information system for this type of transport. The major concepts included in this ontology are: vehicle, route, stop and organization (see Fig. 5).

4. SCHEDULING SYSTEM ONTOLOGIES

The work [Smith, Becker, 1997] includes information about an ontology developed to build scheduling systems. Generally, it concerns production and transport scheduling. This study uses a general scheme resulting from the functional perspective of modelling. The authors extended this schema by additional concepts, as shown in Fig. 6. The exten-

sion consists here in determining the properties of the concepts that are important from the viewpoint of the problem of scheduling.

Demand has, inter alia, the following properties: Product, Earliest-start-date, Latest-start-date, Time-relations (resulting from the constraints imposed by other processes), Priority, Set-of-Activities.

Product has, inter alia, properties as Set-of-Activities, Set-of-Resources.

In the case of the concept of **Resource**, the authors of this study make the list of properties dependent on the type of resource. For this purpose, they include a classification of resources according to three criteria (see: [Smith, Becker, 1997]).

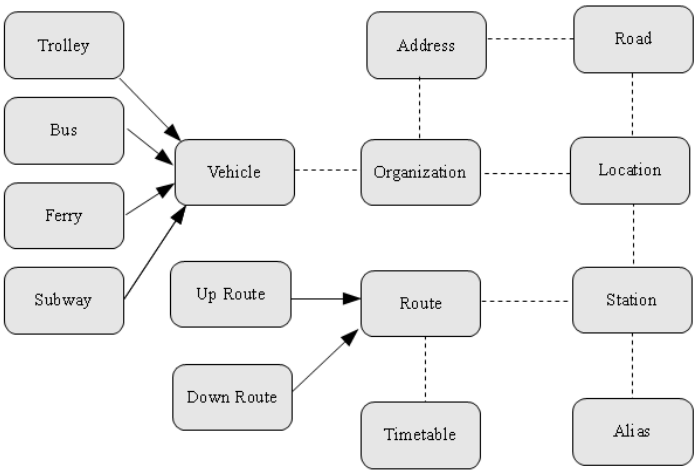

Fig. 5. Diagram of public transport ontology acc. to [Wang, Ding, Jiang, 2005]
Source: Own study based on [Wang, Ding, Jiang, 2005]

Activity in the proposed ontology has the following properties: start time, end time, allocated resources, duration, required resources, relations with other activities, demand, parameters, status.

Constraints in scheduling systems are imposed on the start time, end time and resources. The discussed study presents the following basic types of constraints: value-compatibility constraints, time constraints, resources availability constraints, and instantiation constraints.

The publication [Rajpathak, Mott, Roy 2001], in which concepts related to scheduling have been presented in a more formal way, constitutes particularization of the above work. Axioms associated with the task ontology have also been introduced. The work [Rajpathak et al, (2006)] presents the idea of using ontologies to describe scheduling tasks and the methods for solving these tasks. The authors define a general task of work scheduling as a transformation of nine input variables into an output, which is a schedule (see: Fig. 7). Concepts of ontology representing the input variables consti-

tute the input to the generalized scheduling procedure. The output of the procedures is represented by the concept of Schedule. Diagram of the ontology associated with the scheduling task is shown in Fig. 8. This figure presents the concepts from Fig. 7, as well as other concepts discussed by the authors. Since the ontology is not exactly specified in this work, the contents of this figure should be regarded as a draft.

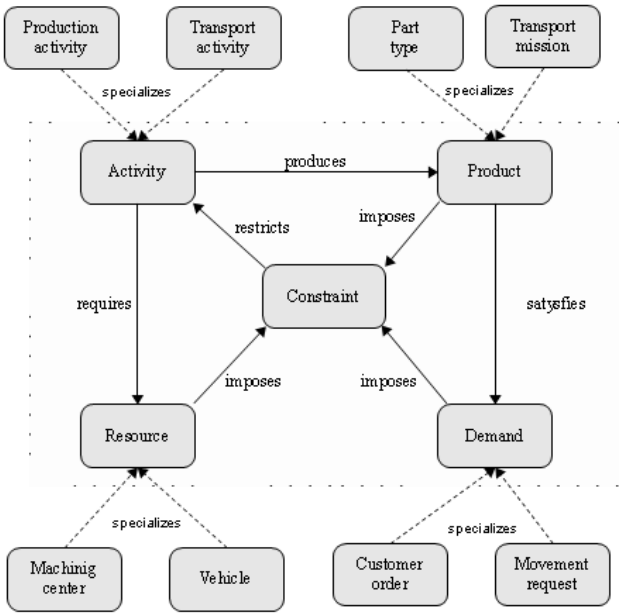

Fig. 6. Basic concepts for the scheduling system
Source: Own study based on [Smith, Becker, 1997]

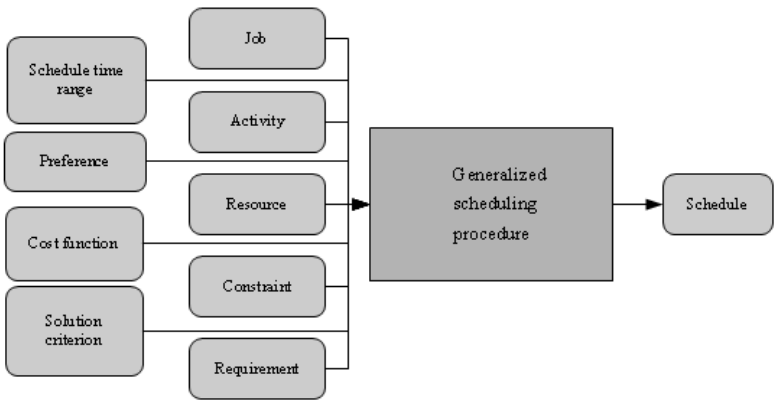

Fig. 7. Context diagram of the generalized scheduling procedure acc. to [Rajpathak et al, 2006]
Source: Own study based on [Rajpathak et al, 2006]

5. SUMMARY

The studies on transportation ontologies were carried out in different directions corresponding to general, domain-specific and task ontologies. First of these ontologies include studies dealing with the fundamental concepts of economics, such as the activity, resource, product and demand. Domain-specific ontologies include the ontologies concerning transport (railway, multi-modal transport), while the task ontologies include those associated with specific tasks, primarily with applications being built. Particularly noteworthy is the ontology concerning the task (problem) of scheduling. The scope of our interest covers various problems related to provision of transport services. These can be classified as strategic, tactical and operational ones.

Decision-making problems, including optimization, take a special place among them. Such a wide range of interests is the reason why the concepts, which regard the business motivation for provision of transport services, demand for carriage, broadly defined transport infrastructure and resources, transport services, as well as problems associated with the provision of transport services, gain importance. A significant part of these concepts has found its place in the available publications.

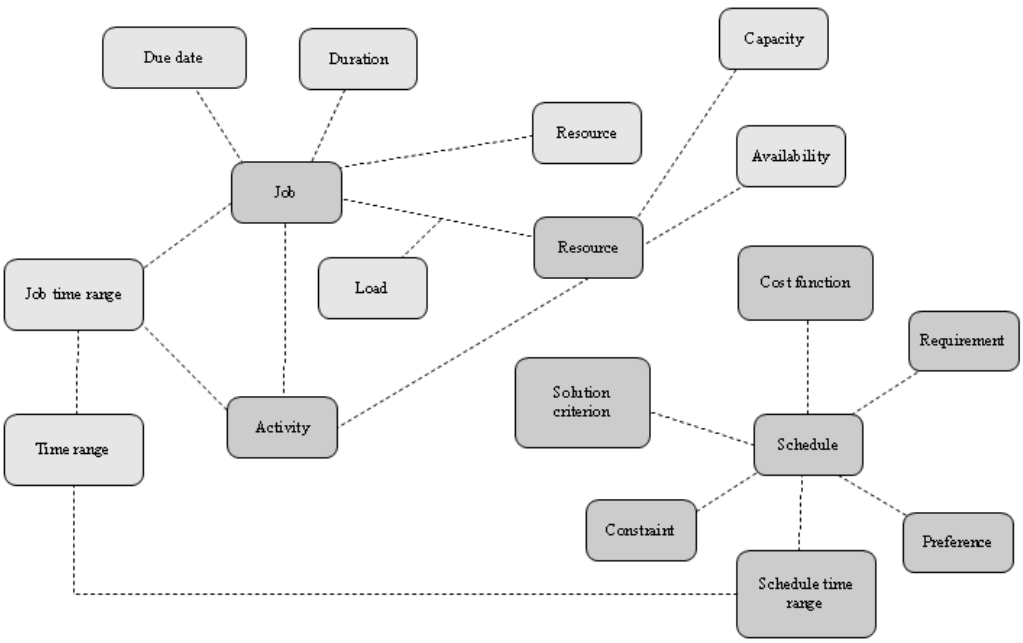

Fig. 8. Diagram of the scheduling task ontology acc. to [Rajpathak et al, 2006]
Source: Own study based on [Rajpathak et al, 2006]

REFERENCES

[2] BECKER M., SMITH S. F. (1997). *An Ontology for Multi-Modal Transportation Planning and Scheduling*. CMU-RI-TR-98-15, The Robotics Institute, Carnegie Mellon University, Pittsburgh, November 1997. (http://www.ri.cmu.edu/pub_files/pub1/becker_marcel_1997_1/ becker_marcel_1997_1.pdf) (06-09-2013)

[3] FOX M.S., GRUNINGER M. (1998). *Enterprise Modelling*, AI Magazine, AAAI Press, Fall.

[4] GRUBER T.R. (1995). *Toward Principles for the Design of Ontologies Used for Knowledge Sharing*. International Journal of Human-Computer Studies, Volume 43, Issues 5–6, November 1995, Pages 907–928.

[5] HOUDA M, KHEMAJ M., OLIVEIRA K., ABED M. (2010) *A public transportation ontology to support user travel planning*. 2010 Fourth International Conference on Research Challenges in Information Science (RCIS), PRINCE Res. Group, Univ. of Sousse, Kairouan, Tunisia, Nice, France, 19-21 May 2010. IBM's Web Services Flow Language (WSFL), 2013, [online], available at: http://www.ebpml.org/wsfl.htm (30-09-2013)

[6] *InteGrail – Publishable Final Activity Report*. Document reference no: IGR-P-DAP-156-07. 13 Apr 2010. http://www.integrail.info/documenti/InteGRail_Final_Project_Report.pdf (03-09-2013)

[7] MARÇAL DE OLIVEIRA K., BACHA F., HOUDA MNASSER, MOURAD ABED. *Transportation ontology definition and application for the content personalization of user interfaces.* Expert Systems with Applications, Volume 40, Issue 8, 15 June 2013 (3145–3159)

[8] RAJPATHAK D. G., MOTTA E., ZDRAHAL Z., ROY R.(2006). *A Generic Library of Problem Solving Methods for Scheduling Applications.* IEEE Transactions on Knowledge and Data Engineering, VOL. 18, NO. 6, June 2006, 815–828.

[9] RAJPATHAK D., MOTTA E., ROY R.(2001). *A Generic Task Ontology for Scheduling Applications.* http://projects.kmi.open.ac.uk/akt/publication-pdf/generictaskontology.pdf (23-09-2013)

[10] SADEGHI-NIARAKI A., KIM K. (2009). *Ontology based personalized route planning system using a multi-criteria decision making approach.* Expert Systems with Applications. Volume 36, Issue 2, Part 1, March 2009 (2250–2259).

[11] SAMPER J.J., TOMAS V.R., MARTINEZ J.J.; VAN DEN BERG L. (2006). *An ontological infrastructure for traveler information systems*. Intelligent Transportation Systems Conference, 2006. ITSC '06. IEEE, Toronto, Ont., 17–20 Sept. 2006 (1197–1202).

[12] SMITH S.F., BECKER M.A. (1997). *An Ontology for Constructing Scheduling Systems*. http://www.aaai.org/Papers/Symposia/Spring/1997/SS-97-06/SS97-06-016.pdf. (03-09-2013)

[13] WANG J, DING Z., JIANG C. (2005). *An Ontology-based Public Transport Query System. Conference on Semantics*, Knowledge and Grid, 2005. SKG '05. First International . Beijing , 27–29 Nov. 2005.

Polish criminal procedure, business process modelling,
queuing networks, discrete-event simulation

Ryszard ANTKIEWICZ, Michał DYK,
Rafał KASPRZYK, Andrzej NAJGEBAUER,
Dariusz PIERZCHAŁA, Zbigniew TARAPATA*

MODELLING, SIMULATION AND COMPUTER SUPPORT OF THE POLISH CRIMINAL PROCEDURE

The Polish criminal procedure is a complex dynamic process created by subprocesses (partially alternative) and atomic activities. Nowadays, the participants of the criminal procedure are provided with modern IT limitedly. Increase the effectiveness of criminal justice is possible via changes in both organizational and procedural way, as well as through the IT technologies. In the paper there is presented IT system supporting tasks of a criminal department of a district court in the two dimensions: (1) processing of information, flow documents and work scheduling; (2) modelling, simulation and optimization of criminal procedure. For the determination of the quantitative characteristics of the criminal procedure's actions, it is proposed to apply models of queuing networks, models of "Activity Network" and "Critical Path" – all of them supported by a dynamic discrete-event simulation.

1. INTRODUCTION

Current analysis and social research on the Polish judiciary indicate negative opinion stated by approximately 40% of the Polish citizens. In particular, the high costs and long delays of criminal cases are criticized. The criminal procedure in its current form and its procedures as well as tools use modern IT technologies very limitedly. Therefore, in this area there is an urgent need for innovation in three directions: technological, organizational and legal.

If we treat the criminal procedure in terms of formal models (e.g. business processes or stochastic processes), the selected problems can be solved using tools for modelling and simulation. Thus, M&S constitutes the basis for optimization, reorgan-

* Faculty of Cybernetics, Military University of Technology, ul. gen. Sylwestra Kaliskiego 2, 00-908 Warszawa 49

ization of the criminal procedure and assessment of both improvement of efficiency and savings (in sense of time and resources). To achieve this objective we proposed the approach based on repeating iterations in cyclical activities: (1) modelling of a courts' structure and organization, the environment and the activities in the criminal procedure; (2) study the quantitative characteristics describing courts and the current criminal procedure; (3) reorganization (optimization) in the direction of improvement; (4) development and implementation of computer systems supporting the courts; (5) re-modelling, study characteristics, verification and correction of direction of the proposed changes and modifications.

In Poland, there are about 330 courts so-called "first instance (district)" courts, almost 50 "second instance (regional)", more than 10 "appellate" and one "supreme (highest)" court. In that structure there are ~10,000 professional judges working in courts. More than ~70% are at first instance, therefore, most of the IT problems belongs to the first instance - the judges need a support in the following areas: managing multiple issues, scheduling of meetings and hearings, as well as create multiple orders and handle hundreds writings.

On the other hand, the "weight" of the considered problem is the opposite - the higher instances (in the district courts or appeal) deal with fewer number of cases, however with a huge amount of information in each case. This means the serious need for support in the field of semantic data processing, data classification and knowledge based on the analyzed data. In summary, the scope of computer-aided process in the criminal departments of district courts (in the field of digital information processing, documents' flow and planning work) is different depending on the instance of a court.

The subject matter is relevant in studies over the world for about 20 years. A great effort has been done to define the performance of judicial systems: the analysis of court work (from the point of view of economics) at [1], statistics or prognosis [2]. The paper presents a new approach resulting from the research based on the case study of District Court in Bialystok.

2. FORMAL MODELLING OF THE CRIMINAL PROCEDURE

Models of cases, activities, people, resources, etc. are defined on the basis of studies carried out in the Criminal Department of the District Court in Bialystok. In order to identify the current modus operandi on the background of different circumstances and surrounded organizations, as well as the target functional model, the "BPM-Business Process Modelling" has been applied. A business process represents ordered sequences of activities from the criminal procedure - composed of other activities and simple (indivisible) - connected by logical operators and initiated by different kinds of events. Models of actions and processes constructed that way are in fact a simplified methodology of judiciary at the district court level.

The Polish criminal procedure, which is strictly based on the Polish Constitution and complex legislation, involves many steps and stages that at the most general level are: Criminal investigation, Grand Jury proceeding and Execution of sentence. Each of these processes is de facto a complex process containing many components of the sub-processes and steps. More detailed, the model describes the following steps: Reporting of the Crime, Pre-arrest Investigation, Arrest, Arrest and Booking Process, Post-arrest Investigation, Decision to Charge, Filing the Complaint, First Appearance, Preliminary Hearing and/or Grand Jury Proceeding, Arraignment on the Information or Indictment, Pre-trial Motions, Plea Bargaining and Guilty Pleas, Trial, Sentencing, Appeals and finally Post-conviction Remedies. The space of steps (and stages of proceeded cases) is perceived as discrete. Depending on goals it can be divided into different number of levels. Number of steps are also optional and may be performed according to various alternative paths - might be terminated at any of the stages as well. The paper focuses on tasks carried out at a stage of criminal proceedings before a court, what is shown in Fig. 1 using BPMN notation.

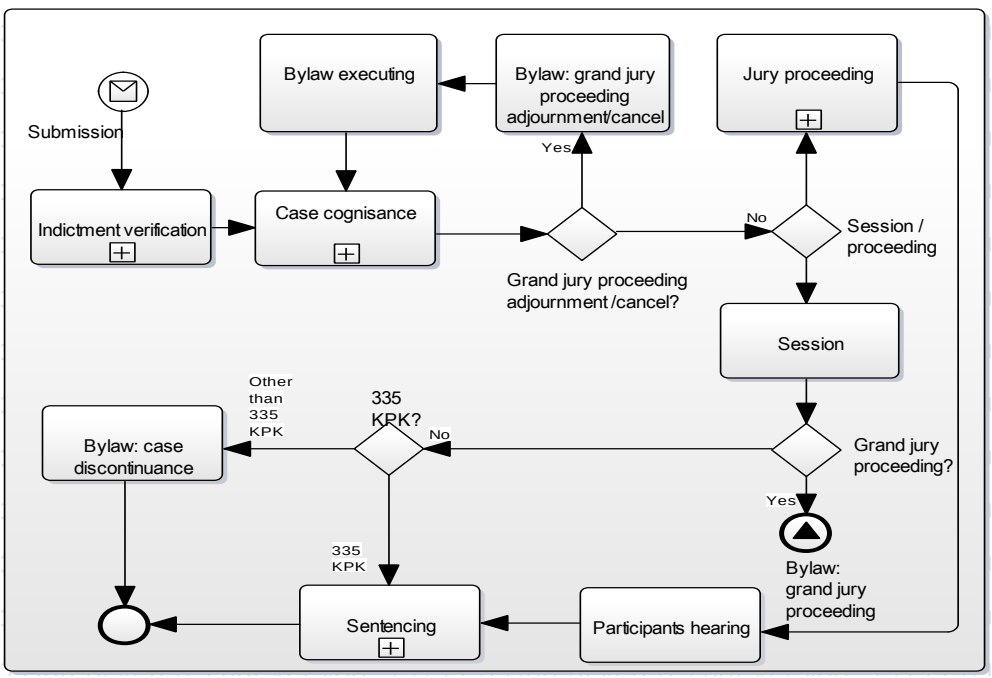

Fig. 1. The diagram of a higher level of judicial criminal proceeding. Source: own preparation

The analysis of procedural activities in a criminal department of a district court has finished with number of such diagrams, which became the basis for the design of computer support system, quantitative analysis and optimization.

3. COMPUTER SUPPORT OF SELECTED PROCESS ACTIVITIES

The conclusion from the analysis of courts' needs and functional deficiencies in the area of IT is that particularly important is to support in the following areas:

- acquisition and flow of digital information to a court and inside a court, including electronic submission office;
- automation of selected steps in the criminal procedure: scheduling meetings, allocation of cases and deadlines, etc.;
- assignment / decreeing / sharing of tasks and employees;
- controlling an order of operations, privileges and conditions of their realization;
- controlled circulation of documents - identification, registration and monitoring of printed judicial documents using QR codes;
- templates of judicial documents - editing, auto-filling with data from the database, validation;
- real time monitoring a status of cases (delay, actions "to do"), outsourced tasks; (e.g. to experts);
- calendars (personal and shared) and a schedule of issues or meetings;
- notifications (message, e-mail or SMS) of essential events, pending tasks or elapsing / overdue dates.

In Fig. 2 the architecture of the proposed system with its surroundings is presented.

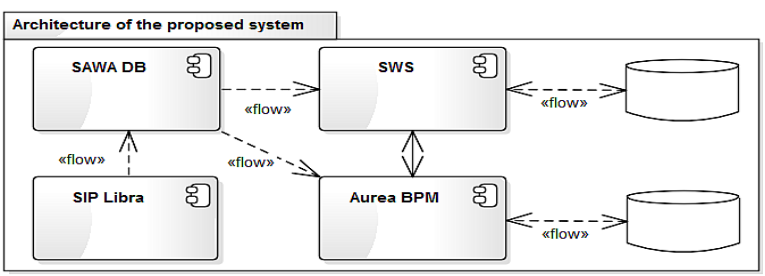

Fig. 2. The proposed supporting system with its surroundings. Source: own preparation

The first module "SWS" is a web application working on the Internet. An access to "SWS" from almost any computer and smartphone devices in a court has different users: department chairmen, judges, employees of secretariats. The input to the system is an electronic submission office through which authorized institutions shall submit documents in electronic form. It launches the criminal procedure outlined in the defined model. In each action at least one document is created or is used. It usually forms the basis for further operations. Analysis of procedural documents (made available by the district court) showed that a large part of the data is repeated in several documents within a single case (e.g. the terms: suspect, accused, convicted might

be correlating with one person). The situation was a demand for an ontology dedicated for criminal procedure. The system uses the component Tecna's "Aurea BPM" to support the correct sequence of steps and automation processing of documents in the editable forms. Each form is associated with a specially prepared document template, MSWord file, a table in a database and a GUI form. In the case of a district court (for jurisdiction actions) there was developed approximately 100 forms.

The subsystem of business processes supports identifying, circulation and tracking of paper documents by marking QR code - every action related to documents is recorded and associated with a user. The subsequent function is monitoring of a status of each criminal case and identifying potential delays. Every judge has view of conducted cases and their status in real time. From this place he has access to electronic form of documents developed during the proceedings. Component "Schedule" supports scheduling of the tasks – it proposes the most convenient dates, taking into account the expected duration and limitations resulting from the already scheduled appointments. The component "Notifications" is strictly associated with "Schedule" - it informs users about changes to deadlines and the elapsed dates.

4. NETWORK MODELS IN THE QUANTITATIVE ANALYSIS OF THE CRIMINAL PROCEDURE

For the determination of the quantitative characteristics of the criminal procedure's actions, it is proposed to apply models of queuing networks, models like "Activity Network" and "Critical Path" – all of them supported by a discrete dynamic simulation if the analytical methods cannot be used.

The first approach - queuing systems model - assumes the realization of the criminal procedure in the form of series of activities assigned to different people. The scope of individual activities, their sequence and method of execution arises from the provisions of the Criminal Code, the Code of Criminal Procedure and the Criminal Executive Code. The possibility of realization of specific activities by specific individuals depends on their respective functions and the findings of a criminal code. Using the defined model we can estimate various quantitative characteristics of the criminal procedure at the stage of judicial proceedings. A precondition is to poses data relating to execution time of each activity and the occurrences of values "Yes" and "No" in the block "Condition" (not only of a single run of criminal case - in practice there exist several concurrent executions). The individuals are requested to perform repeatedly activities in random intervals. Depending on an intensity of documents' submission, some queues with tasks waiting for execution are possible. Therefore, the total execution time of a single criminal case depends both on a time of task's execution and on delay in queue.

To determine the quantitative characteristics of the courts, we propose well-known models and methods of queuing theory and in particular queuing network [7]. The queuing network is defined by means the following vector (1):

$$S = \left\langle \mathbf{W}_0, Q_0, A_0, (B_i, n_i, N_i, f_i)_{i \in \mathbf{W}} \right\rangle \tag{1}$$

where:

- W_0 – set of network nodes' numbers and *zero* node (represents an environment);
- $(B_i, n_i, N_i, f_i)_{i \in \mathbf{W}}$ – the parameters of the internal network nodes;
- A_0 – a type of the input stream (arrival process) of requests (customer, job) incoming into a network;
- Q_0 – a matrix describing a movement of requests inside a network.

In the case of the queuing model of a criminal department we assume that requests (e.g. accusation, customer, job) will be arriving according to Poisson process with intensity λ_0 (the expected number of requests in a fixed time unit). This assumption is justified by the fact that a stream of requests (e.g. accusation) is a result of multiple streams of criminal events occurring in reality. Those events generate the needs of execution of the criminal procedure. In the situation of the criminal procedure at the stage of judicial proceedings the nodes correspond to individuals (or groups) responsible for realization of the procedure's activities.

The model defines different types of requests: $R = \{(r, j) : r = 1,.., R_w,\ j = 1,..L_{cz}\}$, where r is an identifier of a type of the request incoming to a court, R_w is the number of all types of requests (from prosecutors, private prosecutions, ...), j is an identifier of activity of the criminal procedure, and L_{cz} is the number of all activities identified by the model of the criminal procedure. Taking into account the individuals and institutions involved in the execution of the criminal procedure we propose the following nodes (connected like at Fig. 3):

- Node 0 – surroundings – source of accusations and target for the served cases (after a final judgment);
- Node 1 – chairman of a criminal department, the node type of M | G | 1 | ∞ | FIFO with two queues (one contains requests with limited time of service);
- Nodes 2j, $j = 1, .. L_s$ – judges of a criminal department, the node described by M | G | 1 | PS | ∞ | FIFO with multiple queues for different types of requests (in particular with restricted service time); PS indicates that a judge can perform various tasks concurrently, dividing his time between them;
- Node 3 – head of secretariat, the node type of M | G | 1 | ∞ | FIFO;
- Node 4 – secretariat, the node type of M | G | n | ∞ | FIFO;

- Node 5 – public prosecutor (*de facto* it belongs to the surroundings, however during a procedure execution at a judiciary stage it performs operations such as: e.g. replenishment to an accusation, appeal, …) - node type of M | G | ∞ | ;
- Node 6 – private prosecutor - node type of M | G | ∞ |.

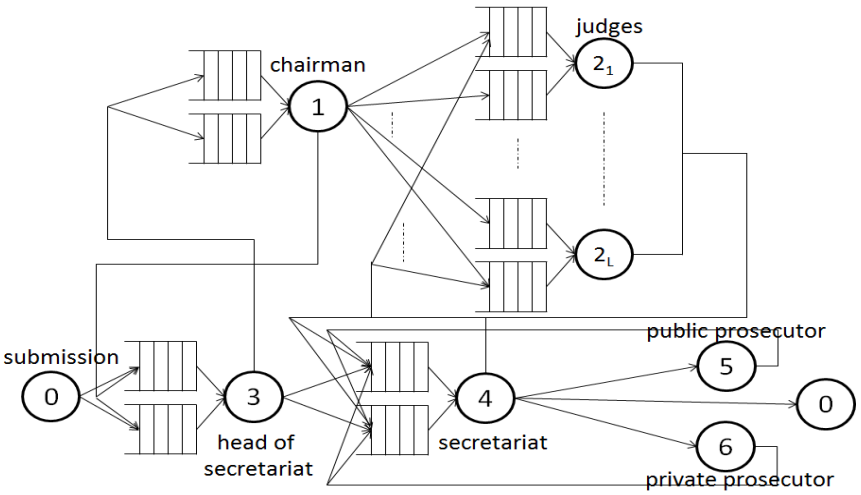

Fig. 3. The queuing model of a criminal department. Source: own preparation

A flow of requests is described by the matrix (2) of transitions' probabilities:

$$Q_0 = \left[q_{i,j}^{(r,k),(r,l)} \right]_{\substack{i,j \in \mathbf{W}_0 \\ (r,k),(r,l) \in \mathbf{R}}} \tag{2}$$

wherein $q_{i,j}^{(r,k),(r,l)}$ means the probability that a request of type *(r, k)*, after service in the node *i-th*, goes to the *j-th* node and becomes the request of type *(r, l)*.
The defined model a criminal department allows the determination of different quantitative characteristics relevant to the assessment of a court's effectiveness:

- distribution and expected degree of utilization of labor resources;
- distribution and expected value of the execution times of the activities performed by individuals or groups;
- distribution and expected value of the execution times of the whole criminal procedures initiated by a specified criminal indictment or application.

Using the characteristics one will be able to identify potential bottlenecks and applied to the optimization process, and thus a work of a court. It should be emphasized that it is impossible to estimate the characteristics of the above-defined queuing network model using analytical methods. However, they can be determined by approximation using BCMP network model [8] or by a dynamic simulation. Construction of an ap-

proximation model relays on the approximation of any distributions of service times for the requests, using phase distributions.

The second approach is the use of computer simulation methods, in particular dynamic and discrete-event [8]. Simulation methods represent the queuing networks in a very adequate way and allow to estimate characteristics at least like analytical methods. At any point in simulation time a new state of the system is determined in accordance with the algorithms that describe behavior of modelled and simulated objects. A simulation event $e_i = < id, t_i, k >$, $id = 1,2, ..., the\ card\ (E)$, $t_i \in T = \{0, 1, 2, ..., n\}$, $k \in K$ of an event class $K = e <^T, f^S >$ (the type of event $e^T \in E^T$ and a function $f^S : S_{i-1} \rightarrow S_i$ of state change) changes the system state discretely. At the moment t_0 system is in initial state $S(t) = S_0$ and the event e_i (occurring in t_i) moves the system from S_{i-1} into S_i. The simulation algorithm changes the global simulation time also in a discrete manner, assigning the time value of a current event: $t^* = min \{t_i: e_i = < id, t_i, k > \}$, $i = 1..2^{ExT}$. Concerning a randomness of activities and events, software generators of random numbers have to be implemented. It conducts towards a multidimensional stochastic process consisting of the events (understood as a state change in the simulation model). In a simulation experiment a number of requests to a court with a given distribution (mostly exponential) of interval is generated and the random path executions of each criminal case are simulated. Every simulated queue is modelled in a form of a group object with FIFO discipline and a theoretically unlimited capacity (in practice: the maximum value a computer can achieved). A server inside a simulated queuing network represents a person (judge, clerk, chairman, secretary) and is modelled as an object characterized by a random service time (a random variable with value set by software random generator of any distribution). Carrying out a series of simulation experiments allows to collect dataset and finally to estimate the statistics on: queues before service nets, partially and completely realized tasks, resource usage, bottlenecks, expected delays or probability of completeness of a case in a given time.

The other proposition considering to determine the characteristics of the criminal procedure are the solutions based on the theories [3], [4], [6]: CPM (Critical Path Method), PERT (Program Evaluation and Review Technique), CPM-cost, PERT-cost, GAN (Generalized Activity Network), GERT (Graphical Evaluation and Review Technique). As an example, consider a part of crime process deals with the first steps after submission of claim form (in Fig. 4).

For "enter" to nodes (events) we have the following logic operations: conjunctive „and", alternative „or", disjunctive „xor" (excluding-or). Consequently, for "exit" from nodes there are logic operations: conjunctive „and", disjunctive „xor" (excluding-or). In the case when all enters and exits for all nodes are conjunctive ("and"-"and") then the network is deterministic (CPM, PERT). Otherwise, the network is stochastic (GERT, GAN). Let's notice that in the criminal procedure at the start and the end of any node in network we can meet each of situations presented above.

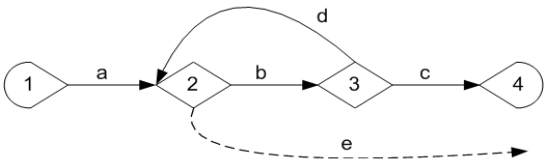

Fig. 4. Example of GERT (GAN) analysis in a part of the crime procedure. Source: on the basis of [5]

There exists the following activities (arcs) in the network: a (submission of claim form - for the first time); b (formal control of claim form); c (transfer of claim form for further processing; d (submission of claim form after removing formal defects); e (equivalent of path from node 2 to node 4 taking into account cycle 2-b-3-d-2). Events (nodes) 1, 2, 3 and 4 describes, respectively: start of the process, event directly after submission claim form, event directly after formal control of claim form and event directly after transferring of claim form for further processing. Nodes 2 and 3 have alternative "enter" and "exit", node 1 has conjunctive "enter" and "alternative" "exit" and node 4 has alternative "enter" and conjunctive "exit" (in this case node 3 should has disjunctive "exit" and node 2 - disjunctive "enter" but for the example of using Elmaghraby's graph algebra the assumption is enough). It causes that the network is stochastic. Let's take into account that the network has the cycle 2-b-3-d-2. This fact - the network (graph) is not acyclic - causes that we can't use classical methods like CPM or PERT. Taking into account Elmaghraby's algebra [4] for GAN/GERT networks we obtain following realization probability and time of the *"e"* activity [5]:

$$P_e = \sum_{n=0}^{\infty} P_a (P_b)^{n+1} (P_d)^n P_c = \frac{P_a P_b P_c}{1 - P_b P_d}$$

$$\bar{\tau}_e = \sum_{n=0}^{\infty} [\tau_a + (n+1)\tau_b + \tau_c + n\tau_d] P_a (P_b)^{n+1} (P_d)^n P_c = P_a P_b P_c \frac{\frac{\tau_a + \tau_b + \tau_c}{1 - P_b P_d} + \frac{\tau_d + P_b P_d}{1 - P_b P_d}}{1 - P_b P_d} \tag{3}$$

where P_x is realization probability of the x activity, τ_x is realization time of the x activity and n is a number (count) of repetitions of the cycle 2-b-3-d-2.

A case when a cycle in a network has more than 2 arcs can be reduced to a case with two arcs only (like in the cycle 2-b-3-d-2) and we can use the procedure described by equation (3). GAN, GERT – "stochastic" events and activities give possibilities to model wider class of problems; possibility of modelling not only "and" enters to nodes and exits from nodes but also "or" and "xor"; problems to set probabilistic distribution of characteristics for events and activities.

4. CONCLUSIONS AND RECOMMENDATIONS

In this work the issue of increasing the efficiency of the criminal procedure with particular attention to the judiciary stage have been analyzed. However, the general goal is elimination of all reasons that might weaken the accessibility of judiciary system to citizens, as well as make hard to keep the guarantee of "equality before the law". To date, the greatest efforts for accelerating criminal proceedings and limiting delays of cases has been directed to increasing the budget of judicial systems – it has been noted in Poland and also Belgium, Luxembourg, Austria. The authors' thesis states that a better method which leads to an increase in efficiency is computer support of selected steps in the criminal procedure followed by operational research. Therefore, we propose two-way analysis of the criminal procedure: modelling and simulation of the process similarly to Business Processes Modelling (BPM) and optimization based on hybrid models and methods: GAN/GERT with queuing theory and optimization. In the case of optimization of a single case (a subprocess inside the whole crime procedure) we obtain locally optimal solutions. In such organizations like judicature (but not only) we must focus on optimizing the operation of whole organization (globally optimal solutions). Thus, particularly important in our study was to model the parallel runs of many criminal cases. The other results are related to identification of "bottlenecks", usage of employee resources and time characteristics of the various stages of the criminal procedure. The computer support system, analytical and simulation tools – have been the case study in the District Court in Bialystok.

Acknowledgments: This work was partially supported by the National Centre for Research and Development: ROB 0021/01/ID21/2 and DOBR/0069/R/ID1/2012/03.

REFERENCES

[1] POSNER R. A., *Judicial behaviour and performance: an economic approach*, Florida State University Law Review, Vol. 32:1259–1279, Florida State University, USA, 2005.

[2] European Commission for the Efficiency of Justice (2012) European judicial systems - Edition 2012. *Efficiency and quality of justice*, ISBN 978-92-871-7559-5, Strasbourg 2012.

[3] EISNER H., *A Generalized Network Approach to the Planning and Scheduling of a Research Project*, Operational Research, Vol. 10 (1) , 1962, 115–125.

[4] ELMAGHRABY S.E., *Algebra for the Analysis of Generalized Activity Networks*, Management Science, 10, 1964, 494–514.

[5] IGNASIAK E., *Teoria grafów i planowanie sieciowe*, PWE, Warszawa, 1982.

[6] LASLO Z., GOLENKO-GINZBURG D., GONIK A., *Alternative stochastic network projects with renewable resources*, Computer Modelling and New Technologies, Vol.9-1, pp. 40–46, 2005.

[7] BOLCH G., GREINER S., DE MEER H., TRIVEDI K.S., *Queueing Networks and Markov Chains*, John Wiley & Sons Inc., Hoboken, New Jersey, 2006.

[8] HARRISON P.G, PATEL N.M, *Performance Modelling of Communication Networks and Computer Architectures*, Addison-Wesley Publishing Company, 1993.

[9] ROSS S.M., *Simulation*, Elsevier Academic Press, 2006.

Flow shop, robust schedule, stable schedule

Dorota KUCHTA*

MACHINE FAILURE ORIENTED ROBUST SCHEDULE IN FLOW SHOP SCHEDULING – A FUZZY APPROACH

A new concept of the robustness of flow machine schedule is proposed, which is based on the features of individual machines. Its aim is to protect the organisation against serious delays caused by accumulation of delays in various processes realised in the organisation at the same time and due to machine failures. A fuzzy model determining the robust schedule defined in the paper is formulated, which can be reduced to a mixed integer linear model. The approach is illustrated by means of an example.

1. INTRODUCTION

Flow shop scheduling has been studied for many years, in various versions (e.g.[2,5]). Also its robustness/stability/sensitivity/security problem has been studied (e.g.[1,2,5,7,9,10,11]). The various problems called schedule robustness/stability/sensitivity/security problems are based on the very true assumption that the job processing times are not always stable and may change with respect to initial estimates.

Without entering into the different robustness/stability/sensitivity/security flow shop problems, let us state that hardly any of them takes into account explicitly two aspects that are of utmost importance in our approach: the type of reasons of the change of processing times and the whole environment of the process in question. Let us look closer at the two aspects:

- The reasons of the processing times being different than initially estimated may be of course of various types. To the knowledge of the author, this variability of reasons has been hardly explored in the literature so far. It is

*Wroclaw University of Technology, Institute of Organisation and Management, dorota.kuchta@pwr.edu.pl

simply assumed that there might be a reason causing a change in the processing times and the processing times are thus regarded as probability distributions or fuzzy numbers. Two exceptions are papers [9,11], which distinguish various types of reasons and point to one type of problems: the machine failure related problems. Thus, our approach aims at protecting the organisation against the consequences of machine failures. We draw inspiration from papers [9,11], but contrary to them we do not use probabilistic approach, but the fuzzy one, which does not require the knowledge of machine failure moments distribution;

- The impact of the prolongation of job processing with respect to the initially planned times may be a serious problem for several reasons – ones linked to the job or the process in question itself (it will be finished later than expected) and ones linked to the environment of the whole process. Here we are interested rather in the latter ones. The processing of each job on each machine may require the usage of resources which are also needed in other processes (material and human resources, machine operators for example) in the given organisation. If a job on a machine is late with respect to the initial schedule (even if it is not the last machine and eventually the whole process or even the job in question will be on time), it may cause quite a nuisance with respect to the coordination all the processes in the organization ([9]). Our approach aims also at protecting the organisation against such problems. Here, we draw our inspiration from the multi-project management.

As a tool to ensure flow shop schedule machine-related robustness, we will use buffers. This term has in the flow shop problem usually another meaning ([8]), but here we will use it in the sense adopted from project management ([3,4]). A buffer will thus be a period of idle time in the schedule, often unknown to the machine operators. It is not explicitly added to the estimated duration time of an activity (job processing), because psychologically thus would push people to use it even if it was not really necessary (the so called student syndrome, [3]). It is an extra reserve, known to the manager of the project or, in our case, process, which gives him more certainty that whole project or process will not be perturbed with respect to the initial schedule and that it will not perturb other processes.

We assume here that each machine has its own characteristics which represent the time needed to repair it once it is broken. This characteristic will determine the size of buffers. The probability or possibility of machine failure may also be taken into account (like in [10]), but here we assume it is the same for all the machines and is considered to be rather high. The impact of various machine failure possibilities (expressed as fuzzy distributions) on the buffer size will be considered in further research. Here we will only allow for the buffers to express (through getting longer and longer) the fact that the failure possibility of a machine may in-

crease with the number of jobs having been processed on the machine so far (the machine will be getting in a way "tired").

The time needed to repair a machine usually cannot be determined in an exact way before the actual failure occurs, as each machine may have various types of failures and people repairing it may have various experience. Also, it will not always be possible/desirable to leave too big buffers, because the buffers will increase the makespan of the process, which is not desirable either (a too long makespan blocks time unnecessarily and makes the organisation lose money). Thus, we apply here the fuzzy approach: the decision maker will give a upper limit for the needed repair time, with which he would be fully satisfied, and the lower limit of his positive satisfaction degree with the repair time. Another decision maker will give a lower bound for the scheduled makespan, with which he would be completely satisfied, and an upper bound, where his satisfaction will reach zero. The satisfaction of the decision maker with the scheduled makespan will be also modeled by means of a fuzzy number.

A mixed integer linear model will be formulated which will determine a schedule maximizing the minimum of all the satisfaction degrees of all the decision makers – the ones interested in the robustness of the schedule with respect to various machines and the ones interested in the scheduled makespan.

2. BASIC INFORMATION ABOUT FUZZY NUMBERS

In this section we refrain ourselves to the type of fuzzy numbers we will need in this paper.

Definition 1 :

1. A left hand fuzzy number $\tilde{A}^l = \left(\underline{a}, \overline{a}\right)^L$ is determined by its membership function $\mu_{\tilde{A}^l}$:$\Re \longrightarrow [0,1]$, where $\Re$ is the set of real numbers, such that there exist two real numbers $\underline{a}$ and $\overline{a}$, $\underline{a} \leq \overline{a}$, such that $\mu_{\tilde{A}^l}(x) = 1$ for $x \leq \underline{a}$, $\mu_{\tilde{A}^l}(x) = 0$ for $x \geq \overline{a}$ and $\mu_{\tilde{A}^l}$ is continuous and decreasing in the interval $\left(\underline{a}, \overline{a}\right)$;
2. A right hand fuzzy number $\tilde{A}^r = \left(\underline{a}, \overline{a}\right)^R$ is determined by its membership function $\mu_{\tilde{A}^r}$:$\Re \longrightarrow [0,1]$, where $\Re$ is the set of real numbers, such that there exist two real numbers $\underline{a}$ and $\overline{a}$, $\underline{a} \leq \overline{a}$, such that $\mu_{\tilde{A}^r}(x) = 0$ for $x \leq \underline{a}$, $\mu_{\tilde{A}^r}(x) = 1$ for $x \geq \overline{a}$ and $\mu_{\tilde{A}^r}$ is continuous and increasing in the interval $\left(\underline{a}, \overline{a}\right)$.

Fuzzy numbers may have various interpretations, but in this paper we will use just one: for a given $x \in \Re$ $\mu_{\tilde{A}^l}(x)$ or $\mu_{\tilde{A}^r}(x)$ represents the satisfaction of a decision maker with value x. Various fuzzy numbers may represent the satisfaction of various decision makers, with various aspects or features or objects, seen from various points of view. Left hand fuzzy numbers will be used in case a decision maker would like a value to be rather small, right hand fuzzy numbers in case big values are preferred. The numbers $\underline{a}$ and $\overline{a}$ represent the limits of satisfaction – beyond the interval $(\underline{a}, \overline{a})$ only full or no satisfaction occurs, in the interval intermediate satisfaction degrees are contained.

3. FORMULATION OF THE BASIC FLOW SHOP PROBLEM

We are dealing with the basic form of the flow shop problem: we have *m* machines and *n* jobs. p_{ik} $(i = 1, \dots, n, k = 1, \dots, m)$ stands for the processing time of the *i*-th job on the *k*-th machine. Index $i = 1, \dots, n$ refers to the initial ordering of the jobs. Each job has to be processed on each machine in the sequence from the 1st machine through the 2nd etc. till the *m*-th one. A function p: $\{1, \dots . n\} \rightarrow \{1, \dots . n\}$ is searched for such that if the elements are processed in the sequence determined by $p(i)$, the makespan of the process will be minimal. The order given by function *p* will be denoted using index *j*.

If C_{jk} (j $= 1, \dots, n, k = 1, \dots, m$) stands for the finish moment of the processing of the *j*-th job on the *k*-th machine, the makespan of the whole process is equal to C_{nm} and it should be, in the basic formulation of the problem, minimized.

The following mixed integer model for the basic flow shop problem is based on the formulation from [6,10]. Following decision variables (apart from C_{jk} $(j = 1, \dots, n, k = 1, \dots, m)$) are used:

- x_{ij} – equal to 1 if t $p(i) = j$ and 0 otherwise, $i = 1, \dots, n;\ j = 1, \dots, n$
- I_{jk} – the idle time of the *k*-th machine between the processing of the *j*-th and (*j*+1)-st job, $j = 1, \dots, n-1, k = 1, \dots, m$;
- W_{jk} – the waiting time of the *j*-th job between the *k*-th and the (*k*+1)-st machine, $j = 1, \dots, n, k = 1, \dots, m-1$.

The model itself is as follows:

$$C_{nm} \rightarrow min \tag{1}$$

$$C_{1m} = \sum_{k=1}^{m-1}(\sum_{i=1}^{n} x_{i1}p_{ik} + W_{1k}) + \sum_{i=1}^{n} x_{i1}p_{im}; \tag{2}$$

$$C_{jm} = C_{j-1,m} + I_{j-1,m} + \sum_{i=1}^{n} x_{ij}p_{jm},\ j = 2, \dots, n; \tag{3}$$

$$I_{j,k} + \sum_{i=1}^{n} x_{i,j+1}p_{ik} + W_{j+1,k} = W_{j,k} + \sum_{i=1}^{n} x_{ij}p_{i,k+1} + I_{j,k+1},\ j = 1, \dots, n-1, k = 1, \dots, m-1; \tag{4}$$

$$\sum_{j=1}^{n} x_{ij} = 1 \text{ for all } i = 1, \dots, n \quad (5)$$

$$\sum_{i=1}^{n} x_{ij} = 1 \text{ for all } j = 1, \dots, n \quad (6)$$

Constraints (2) and (3) are involved in obtaining the completion times of each job on the *m*-th machine. Constraint (2) applies only to the job ordered in the first position, whose completion time depends on its own processing times and the waiting time between the machines. Constraint (3) is a recursive constraint, referring to all the other jobs. Constraint (4) expresses the relationship among processing, waiting and idle times - details with an explanatory illustration can be found in [6,10]. Constraints (4) and (5) ensure that a job can be allocated only to one sequence position and that each sequence position can be related only to one job.

Once the solution is determined, the values of the decision variables x_{ij}, i $= 1, \dots, n$; $j = 1, \dots, n$ will give the function *p*, thus the optimal sequence.

Let us illustrate model (1)-(5) with an example:

Example 1: Let us consider two machines and three jobs. The processing times are as follows:

Table 1. Data for Example 1

Job i	p_{i1}	p_{i2}
Job 1	2	7
Job 2	3	5
Job 3	4	6

In this case we get: x_{11}=x_{32}=x_{23}=1 , thus $p(1) = 2, p(3) = 2, p(2) = 3$ and the optimal order is: Job 1 , Job 3, Job 2 (this is not the only optimal solution, but we are not interested here in the problem of alternative solutions). The optimal value of C_{32}=20, and we get the values C_{12}=9 and C_{22}=15, I_{11}=I_{12}=I_{21}= I_{22}=0, $W_{11} = 0$, $W_{21} = 3$, $W_{31} = 6$, which corresponds to the following schedule:

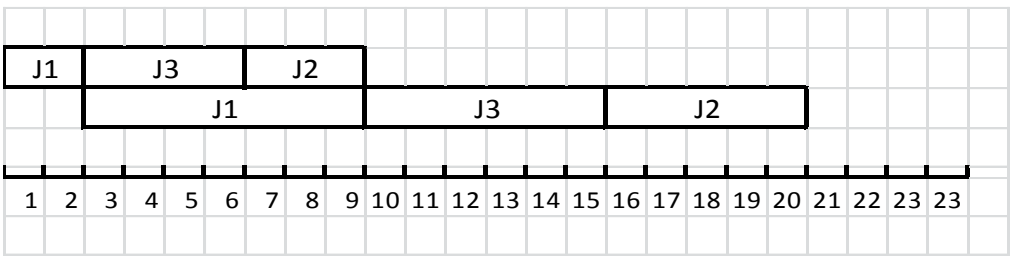

Fig. 1. Optimal schedule for Example 1

In the schedule from Fig. 1 there may be a delay on machine 1 while processing Job 2 and Job 3 without affecting the optimal makespan 20, but any other type of delay would change the makespan. We can see that this schedule is more "robust" (we are using this term here without an exact definition yet) with respect to the first machine than with respect to the second. It would be easy to formulate examples where the schedule would be more robust with respect to the second machine than with respect to the first. Now, our aim is to ensure robustness with respect to both machines. This subject will be made concrete in the following section.

4. FORMULATION OF THE MACHINE-RELATED ROBUST FLOW SHOP PROBLEM

If we choose schedule from Fig.1, we have to reserve resource (e.g. machine operators, potentially different, which know both the specific machine and the specific job) for the following periods: 1-2 (Job 1 on the 1. machine), 3-9 (Job 1 on the 2. machine), 3-6 (Job 3 on the 1. machine), 10-15 (Job 3 on the 2. machine), 7-9 (Job 2 on the 1. machine), 16-20 (Job 2 on the 2. machine). If in any of those periods something goes wrong, and here, in this paper we concentrate above all on problems related to the machines, there might be a problem, as the corresponding piece of resource outside the reserved periods may be needed in another process and the delay may not only influence the process in question, but many other processes in the organization.

The slacks mentioned in the previous section, not affecting the makespan, might allow us to use the schedule illustrated by Fig.2 and Table 1:

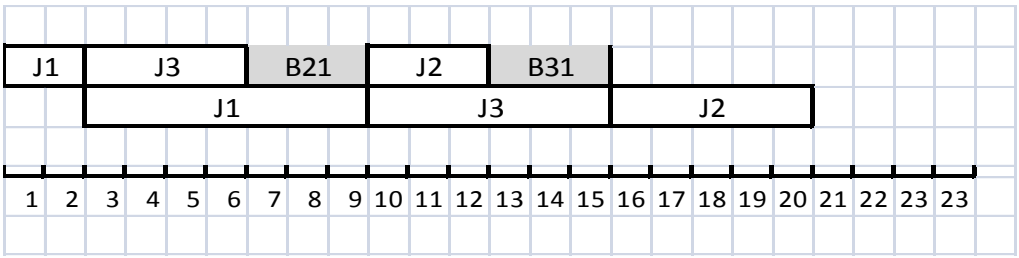

Fig. 2. Tentative usage of buffers (not affecting the makespan) in Example 1

Notation of the buffers, *Bjk*, is analogous to that of the idle times and the waiting times defined earlier (I_{jk} and W_{jk}), we change the name *B* ("buffers") to underline the role of these periods in the multi process management in the organisation we want to stress here, the role analogous to the role of buffers played in project management.

Thus we might reserve the resources also for the buffers (see Table 1), so that in case of a machine failure possibly no more serious, no global problem occurs:

Table 2. Tentative usage of buffers in Example 1.

Resources needed for realizing:	know they have to do the job in the period:	are reserved (by the higher management) for the period:
Job 1 on machine 1	1-2	1-2
Job 1 on machine 2	3-9	3-9
Job 2 on machine 1	**10-12**	**10-15**
Job 2 on machine 2	16-20	16-20
Job 3 on machine 1	**3-6**	**3-9**
Job 3 on machine 2	10-15	10-15

In Table 1 bold rows show the usage of buffers which are in favour of the objective set in this paper: even if machine 1 fails during the realization of the 2. and 3. job , there will be time to repair it and the resources will be free to wait, because, thanks to the buffers, they will be not assigned to any key operation elsewhere in the organisation. At the same time, they will not be aware of the existence/size of the buffers, so that they will not fall prey to the student syndrome.

However, it is clear that the schedule from Fig.2 is not completely robust: if a failure occurs during the realisation of Job 1, the schedule has no protection, and it is completely unprotected against failures of machine 2. The model we will propose now will determine a schedule which will be to some extent protected with respect to all the machines. But this cannot be achieved without sacrificing the makespan, at least to some degree.

We assume that the time needed to repair the machines is not crisp and can be given by experts. We will let them define the levels of the best and the worst accepted protection level for each machine. It will be expressed by means of right hand fuzzy numbers in the following way: $\widetilde{M}^r{}_{jk} = \left(\underline{m}_{jk}, \overline{m}_{jk}\right)^R$, $k = 1, \dots, m, j = 1, \dots, n-1$ will represent the satisfaction of the decision maker responsible for the k-th machine with the length of buffers left between the scheduled processing end and start of the j-th and the (j+1)-st consecutive jobs realized on the k-th machine. Here, it is possible to take into account the "tiring" of the machines and increase $\underline{m}_{jk}$ and/or $\overline{m}_{jk}$ with j, j=1,..,n-1, for a given k, k=1,...,m, if this is appropriate for the given situation. At the same time, a left hand fuzzy number $\tilde{S}^l = \left(\underline{s}, \overline{s}\right)^L$ will represent the satisfaction of another decision maker with the scheduled makespan of the process. A machine-failure oriented robust optimal schedule will be then defined as follows:

Definition 2: For given $\widetilde{M}^r{}_{jk}$, $k = 1, \dots, m, \mathrm{j} = 1, \dots, n$ and $\tilde{S}^l$, other data being defined as in model (1) – (5), the machine-failure oriented robust optimal schedule is one fulfilling constraints (2) – (5) with the following objective function:

$$max\left(\left\{\left\{min\left\{\mu_{\widetilde{M}^r{}_{jk}}(I_{jk}),\mu_{\widetilde{M}^r{}_{jk}}(W_{jk})\right\}_{j=1}^{n}\right\}_{k=1}^{m},\mu_{\tilde{S}^l}(C_{nm})\right)\right. \tag{6}$$

Thus, we strive to have the idle times satisfying the machine-related or prudence-oriented decision makers to a possibly high extent, but we do not forget the decision maker who is interested in a short scheduled makespan.

A model which would allow us to find the machine-failure related optimal schedule is one comprised of objective function (6) and constraints (2) – (5). However, this model would not be linear. If we assume that the decreasing and the increasing parts of membership functions in Definition 1 are linear, it can be shown that we get obtain a mixed-integer linear programming model which would permit us to find a machine-failure related robust optimal schedule.

Let us illustrate the approach for Example 1. We assume: the decision maker responsible for the maintenance of machine 1 and 2 would be fully satisfied if the idle (waiting) time on the machine after each job would be 5, the lower limit for his satisfaction is 0, the machine tiring is to be neglected on both machines, the decision maker responsible for the makespan says he would be fully satisfied if the makespan was 20 (the optimal value), and the limit of his satisfaction is 40. The solution of the respective model is presented on Fig.3:

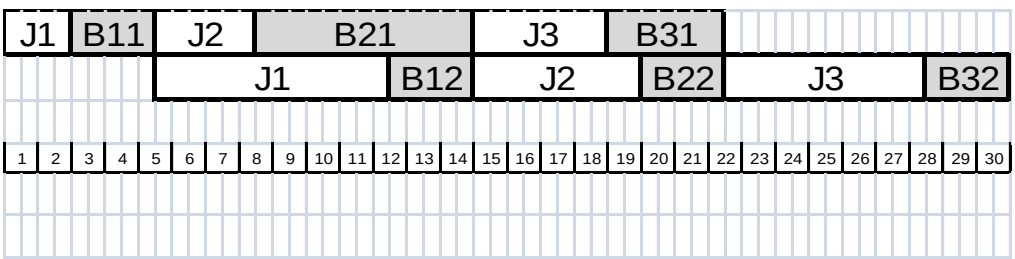

Fig. 3. The machine failure related optimal schedule for Example 1

The value of the objective function (6) is 0,5, and it is attained on buffer B11 on the first machine, on buffers B12 and B22 on the second machine and on the makespan, equal to 30. The other two buffers on the first machine give the decision maker a higher satisfaction degree.

The reservation of resources, assuring the robustness of the schedule, would be as follows:

Table 3**.** Usage of buffers in Example 1, robust schedule.

Resources needed for realizing:	know they have to do the job in the period:	are reserved (by the higher management) for the period:
Job 1 on machine 1	1-2	1-4,5
Job 1 on machine 2	4,5-11,5	4,5-14
Job 2 on machine 1	4,5-7,5	4,5-14
Job 2 on machine 2	14-19	14-21,5
Job 3 on machine 1	14-18	14-21,5
Job 3 on machine 2	21,5-27,5	21,5-30

CONCLUSIONS

In this paper we propose a concept of a machine failure related robustness of the flow shop problem. It takes into account the time each machine may need to get repaired once it is broken. The buffers inserted into the schedule permit to protect the whole organisation against problems due to machine failures in various processes executed in the organisation. Fuzzy modeling is used to express the comprise of the need for robust and for short schedules, but finally a mixed integer linear programming problem is enough to express the problem and solve it.

REFERENCES

[1] ALOULOUA M.A., ARTIGUES CH. (2010), Flexible solutions in disjunctive scheduling: General formulation and study of the flow-shop case, Computers & Operations Research 37, p.890–898;

[2] BAKER K.R. TRIETSCH D. (2009), Principles of Sequencing and Scheduling, Wiley and Sons, p.137–164;

[3] GOLDRATT E. (1997), Critical Chain, North River Press, Great Barrington;

[4] HERROELEN W.S., LEUS R. (2004), The construction of stable project baseline schedules, European Journal of Operational Research, 156(3), p. 550-565;

[5] JENSEN M.T. (2001), Improving robustness and flexibility of tardiness and total flow-time job shops using robustness measures, Applied Soft Computing 1, p.35–52

[6] KANET J.J., SRIDHARAN V. (2000), Scheduling with inserted idle time: problem taxonomy and Operations Research 48, pp. 99–110;

[7] KASPERSKI A., KURPISZ A., ZIELIŃSKI P. (2012) ,Approximating a two-machine flow shop scheduling under discrete scenario uncertainty, European Journal of Operational Research 217, p.36–43;

[8] LEISTEN R. (1990), Flowshop sequencing with limited buffer storage, International Journal of Production Research 28, pp. 2085–2100;

[9] SABUNCUOGLU, I. GOREN, S., (2009), Hedging production schedules against uncertainty in manufacturing environment with a review of robustness and stability research. International Journal of Computer Integrated Manufacturing, 22(2), p. 138–157.

[10] STAFFORD E.F. (1988) , On the development of a mixed-integer linear programming model for the flowshop sequencing problem, Journal of the Operational Research Society 39, p. 1163–1174;
[11] XIONG J., XING L., CHEN Y. (2013), Robust scheduling for multi-objective flexible job-shop problems with random machine breakdown, Int. J. Production Economics 141, p. 112–126.

ACKNOWLEDGEMENTS

This work is supported by Polish National Science Centre grant "The buffer size and location in a proactive and reactive risk management of a project duration over-running" 2011/01/B/HS4/02821.

Customer relationship, customer knowledge,
customer relationship management, knowledge management,
interactive marketing, event based marketing, real time marketing,
advanced analytical platform for knowledge
customer relationship management

Jarosław A. CHUDZIAK,
Andrzej CIEMSKI*

ACQUISITION, GENERATION AND MANAGEMENT OF THE CUSTOMER KNOWLEDGE IN AN INTERACTIVE MARKETING

This paper presents an approach and experiences in building business and technology environment for acquisition, generation and management of the customer knowledge in interactive marketing model. It discusses maturity models of current marketing including event-based marketing and real-time marketing. Then it shows a customer insight process from acquisition up to analytical models which provide proper, treatment-oriented segmentation with propensity models. It lists required functionality of various components of technology which enable creation of platforms for customer knowledge driven customer relationship management. It describes also use of mathematical models and other necessary tools to build analytical solution for continuously increasing demand for adequate data from interactive marketing. In addition, short discussion regarding practical implementation is attached with special comments on efficient and effective implementation of the proposed solution.

1. ROLE OF INTERACTIVE MARKETING IN CURRENT MARKETING ACTIVITIES

1.1. DEFINITION OF INTERACTIVE MARKETING

The modern interactive marketing has started to be one of the dominant approaches in building the organization strategy[1], the current academic studies and practical

* Institute of Computer Science, Warsaw University of Technology

[1] CHUDZIAK J. A., CIEMSKI A., *Implementation of innovative methods of customer relationship management for marketing strategies using analytical tools, accepted for publication, 1st* National Scien-

solutions in various industries. The speed of technology innovation diffusion in many aspects of people interaction in various channels[2] makes solution more and more mature and ready for sustainable solutions.

The interactive marketing is about how to properly use various interactions in all possible channels of communication (mobile, internet, call center, POS, external agents etc.) with existing or potential customer to engage them and then, develop, manage, retain etc. It covers campaign management based on identified events related to customer behaviors, brand communication management etc.

One of the key enabler and prerequisite for well-defined interactive marketing is proper implementation knowledge management framework for customer data. It creates complex ecosphere which should cover knowledge definition (client analytic record, 360 degree view, analytical models), knowledge acquisition from internal and external sources (small and big data, static and dynamic, often with forecasted data with propensity values), plus effective search mechanisms, simulations platforms for tactical and operational problems (pricing, campaign optimization, distribution channel optimization), trends monitoring (evolutions of segments definitions, migration of customers), rule-based reasoning for various part of the solution (nest best action, particular lead optimization).

Another dimension very often considered is real-time solution. It requires adequate integration of interaction channels with our knowledge base customer management system in particularly with the effective and efficient analytical engine which can provide adequate decision enabling recommendations to real or artificial agents or interacting avatars.

The business components of interactive marketing solution contains: product factory, client segmentation, customer journey analysis and then design, marketing and sales events identification and/or creation, events monitoring and management, lead generation rules, campaign optimization policy, channel management, client experience definition, client experience monitoring.

1.2. MATURITY LEVELS OF MARKETING SOLUTIONS

The marketing is undergoing big evolution due significant technology change which influences customer life and buying behaviors. Let's discuss briefly three levels of marketing approaches which currently exist in practice: traditional product-centric, event-based and its extended version real-time.

---

tific Conference, Modern problems of management, organization and production engineering, Cieszyn, Poland, 30 May, 2014

[2] BLATTBERG R. C., KIM B-D. NESLIN S.A, *Database Marketing: Analyzing and Managing Customers*, Springer 2008

Traditional marketing operates basing on the product centric model. The logic of this approach assumes at the beginning of marketing activities availability of the product. Then the effort is placed to find suitable customers for specific product. Uniform, consistent product is communicated to customers in a mass manner. The communication is done by marketing campaigns targeted to the large number of customers using outbound channels. The approach features a weakness in addressing the current customer's product needs and a low marketing efficiency rate.

The next level of maturity in the development of approaches used in marketing activities is the event based marketing. The logic of this approach is opposite to the logic used in the product centric model. In the center of marketing activities is placed a customer. Having a customer the marketing effort is related to discover customer products needs finishing with the proposition of the personalized offer delivered at the right time. These needs are recognized by identification specific customer behavior events. Making marketing communication the product is tailored to the specific customer in one to one manner and is delivered by preferred customer outbound channels. This approach features the higher marketing efficiency rate than achieved by traditional approach.

In real time marketing every inbound customer contact is treated as a potential sales lead contact. The logic of this approach is the same as used in the event based marketing: the customer centric approach, right channel and right time. The real time marketing uses both inbound and outbound communication channels. Additionally a tailored offer is proposed based on the customer needs and in the context of interaction with customer. This approach features also the higher efficiency rate than achieved by traditional marketing approach. The customer centric model is a starting point for an initiation process for an introduction profitable marketing programs[3].

2. KNOWLEDGE IN THE CUSTOMER CONTACTS MANAGEMENT

2.1. CUSTOMER INSIGHT - THE KNOWLEDGE ABOUT THE CUSTOMER

Knowledge of the customer should be consistently built in a systematic way. Each customer behavior, customer activity, customer contact create the possibility to get data, information and knowledge about the customer.

The contact, the relationship with the customer and the use of customer knowledge during the contact inspired scientists to build new disciplines in the in the management and information theory. In the theory of management a new discipline has been

[3] HUGES A. M., *Strategic Database Marketing 4e: The Masterplan for Starting and Managing a Profitable, Customer-Based Marketing Program*, Mc-Grow Hill 2012

formulated called customer knowledge management[4]. The argument has been also stated here that relationship with the customer can be improved by knowledge application[5]. In the area of the information technology a new type of information systems has been defined and build called knowledge customer relationship management (KCRM)[6]. The basic concepts used in the customer contact management supported by customer knowledge became: high value customers and customer's lifecycle.

Customer knowledge management is the process with three general elements: knowledge acquisition, generation and management. Below we list these elements with relevant description and correlation with customer analytics.

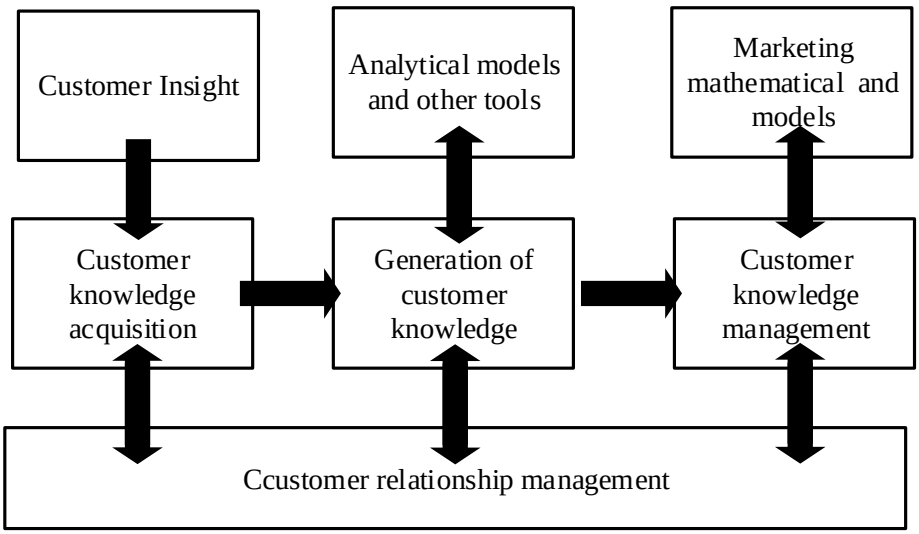

Fig. 1. The correlation between customer relationship and knowledge
Source: Own elaboration

2.2. CUSTOMER KNWOLEDGE ACQUISITION

In the order to better understand the customer behavior the approach based on the insight can be applied[7]. The first step in the interactive marketing is the prop-

[4] BOLJANOWIĆ J. A., STANKOVIĆ J., *The role of knowledge management in building client relationship*, Singidunum Journal of Applied Science 2012: 16-28

[5] WILDE S., Customer Knowledge Management Improving Customer Relationship through Knowledge Application, Berlin, Springer-Verlag, 2011

[6] TIWANA A., *The essential guide to Knowledge Management E-business and CRM Applications*, Prentice Hall, 2000

[7] FOREMAN J. W., *DATA SMART - using data science to transform information into insight*, John Wiley & Sons 2014

er definition of data sources for customer insights. These sources of data can be a base for quantitative insight and qualitative insight. The quantitative insight deals with the numbers such as age, income, average balance of the account, last purchase, demographic etc. The sources of data for quantitative insight can be operational data stores e.g. transactional data, customer data, product data. The qualitative insight are coming from the customer observation and contacts e.g. ability to take risk investment, brand recognition. From the knowledge perspective qualitative insights may take form of the explicit and tacit knowledge. The sources for explicit knowledge can be e.g. notes making by employees during the contacts with the customers, analysis social media content, customer questionnaire. The sources for tacit knowledge can be e.g. customer behavior observations. This kind of the knowledge is gathering in the intuitive way be specialists and experts during their work.

2.3. GENERATION OF CUSTOMER KNOWLEDGE

In the interactive marketing knowledge acquisition is the first step in customer knowledge management. The acquired knowledge should be developed and transformed into a structured form.

A key element in this step is the division of customer population according to their values, needs and attitudes. The division of customer population according to such criteria is very complicated organization activity at the conceptual and computational level. This organization activity can be supported by analytical models. The division of customers is done by using clustering methods which provide to grouping customers into similar segments. Based on the customer segmentation detailed description of customers is done in the form of customer profiling. Customer profiling describes in detail current customer value, product, needs, attitudes, preferred channel contact etc.

2.4. MANAGEMENT OF CUSTOMER KNOWLEDGE

A key element in this step is the proper customer knowledge sharing, distribution and management. The gathered structured and unstructured customer knowledge can be share in the form of customer repository. The quality of the knowledge gathered can be measured by e.g. monitoring trends e.g. evolution in segment definitions, segment value change, migration of customers. The valuation methods from mathematical marketing can be implemented here.

3. PLATFORM FOR KNOWLEDGE CUSTOMER RELATIONSHIP MANAGEMENT

3.1. COMPONENTS OF THE PROPOSED SOLUTION

The key element for integration customer knowledge relationship management and event/real based marketing is technology which should provide solutions to work in real time (close to real time) during the process of acquisition, generation and management of customer knowledge. The platform using such technology enables in the effective way retain current customers, obtain new customers by provided for them tailored offer at the right time.

In the solution we distinguish the following components: customer knowledge acquisition, customer knowledge generation, event and real time marketing, customer contact and lead management and finally knowledge customer management.

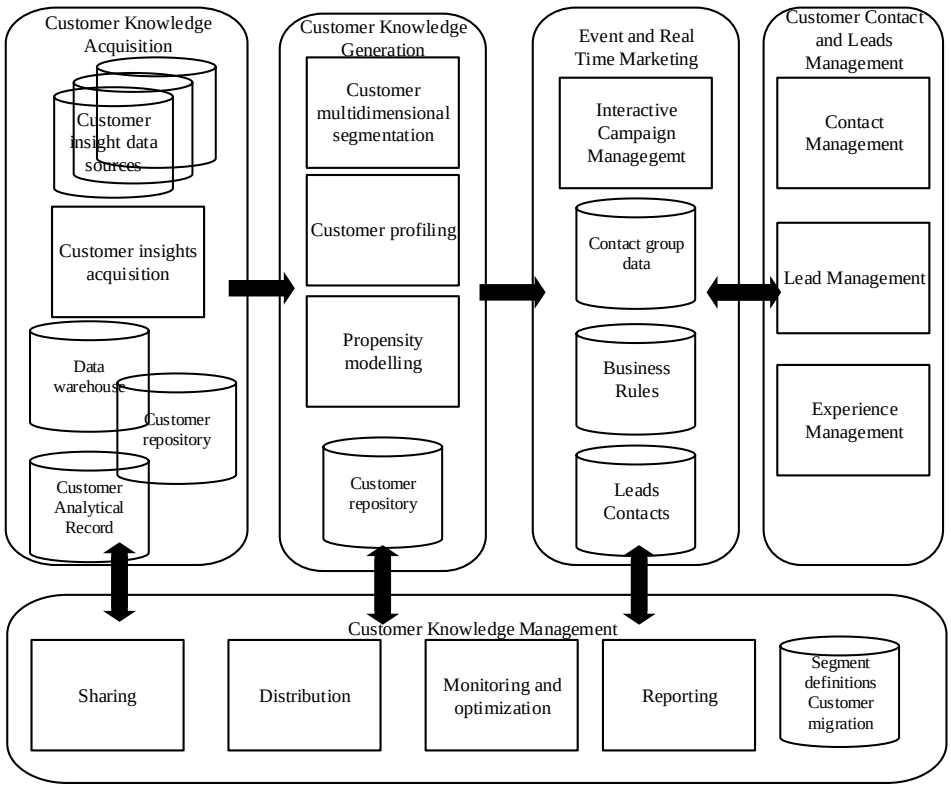

Fig. 2. Knowledge customer relationship management platform for interactive marketing
Source: Own elaboration

Customer knowledge acquisition component consists of defined data sources for customer quantitative and qualitative insights e.g. transactional data sources, social media data, internet and mobile data, intranet data, chats, survey forms, customer data experiences etc. The component uses the functionality (extract, transform, load, document management, collaborative, workflow etc.) called customer insights acquisition to divide and collect data properly in the form of data warehouse and customer repository. The data warehouse and customer repository are the base for further customer insights and knowledge generation: e.g. customer analytics, customer segmentation, customer profiling, data mining and artificial intelligence.

Customer knowledge generation component is responsible mainly for customer segmentation and profiling. Additionally component uses propensity modeling functionality to prepare treatment-oriented customer segmentation (predictive segmentation). Propensity modeling can use e.g. data mining techniques (predictive modelling - decision trees[8], regression, neural networks) and propensity score methods.

Customer knowledge management component is responsible for sharing and distribution customer knowledge. The proper identification of sharing and distribution of knowledge between KCRM and event/real marketing should be defined. The component uses monitoring and optimization functionalities to assess the quality of gathered customer knowledge for interactive marketing e.g. segment value changes, migration of customers.

To specify such advanced platform the project methodology should covers specific approaches to process specification, construction, deployment, testing and monitoring performance of the solution[9].

3.2. THE LOCATION OF THE INTERACTIVE MARKETING ON THE KCRM PLATFORM

The entire gathered customer knowledge (starting from customer qualitative and quantitative insight process) in the form of customer analytical record, customer segmentation, profiling, propensity outcomes supported by expert knowledge (explicit and tacit) is used to build contact data groups with their priorities, business rules for offer selection with their priorities and finally offer with their priorities. The population of customers can be divided according to contact data groups. The different divisions of customer population create target groups for marketing campaigns.

---

[8] SZCZERBA M, CIEMSKI A., *Credit Risk Handling in Telecommunication Sector*, Advances in Data Mining, Applications and Theoretical Aspects, 9th Industrial Conference, ICDM 2009, Leipzig, Germany, July, 2009. Proceedings. Lecture Notes in Computer Science Springer 2009

[9] MODRZEJEWSKI, J. A. CHUDZIAK, R. W. CEGIELSKI, *Complex Marketing Database Specification, Design and Implementation, Computer Information Systems and Industrial Management Applications*, 2008. CISIM '08. 7th

Contact group data consist of data from different customer knowledge areas: customer value, future customer value, customer propensity to purchase, preferred contact channel, currently available products, desirable products, current marketing campaign etc. Business rules for offer selection are built based on the contact data groups.

Interactive campaign component is responsible for providing the appropriate marketing communication in the form of the optimized lead i.e. particular campaign, customer, offer and right channel. In the event/real based marketing campaigns marketing distribution must be adopted to customer interactions and events. The dialogue with customer is being initiated. The flow of the dialogue depends on the feedback from the client.

4. USE OF MATHEMATICAL MODELS AND OTHER NECESSARY TOOLS TO EVALUATE SOLUTION FOR ITERACTIVE MARKETING

4.1. MATHEMATICAL APPROACHES AND MODELS FOR INTERACTIVE MARKETING

Building effective and optimized components of the customer knowledge framework (acquisition, extraction, management) for interactive marketing requires properly usage models and analytical tools. Below we lists the need of use these approaches, models and tools.

As it has been stated before the base for knowledge generation is customer analytical record and multidimensional customer segmentation. The key issue in the process of building customer analytical record is the proper selection of customer variables. The selection of a particular variable can be made by expert knowledge judgment about variable significance. Formally, data quality assurance methods can be applied by using the base statistical analysis, analyzing null and zero values and the variety of variables with agreed cut-off levels for value reject decisions. For specific values of variables (outliers, null, negative) mathematical transformation can be used in order to achieve required domains and scales for those values. In order to build additional analytical variables for data mining purposes different aggregation functions can be used. According to semantic and business goal of the variable/variables different aggregation functions can be used e.g. average, minimum, maximum, count, first value, last value, logarithm.

Having the full 360 degree of customer view and taking into account all possible valuable data insight (qualitative insight, quantitative insight, insight from small and big data, insight from social media etc.) checked against the availability, completeness and quality the process of multidimensional customer segmentation can be started.

For building customer segmentation un-supervising methods of machine learning can be applied. For instance clustering methods are used here with the e.g. centroid algorithm and a Euclidian distance as a measure between objects. The outcome of this

algorithm is the number of similar groups of data. To limit the number of variables used by clustering algorithm, methods of reduction dimensions are used to choose those variables which explain the most variability in the whole data. For testing and evaluating the clustering solution the whole population is dived to training, validating and test data. Finally the profiling of customer groups of data is prepared with using approach based on average in the cluster with the comparison to average in the whole population. The description (profiling) is done for each important variable.

4.2. MONITORING THE EFFECTIVENESS OF INTERACTIVE MARKETING

The effectiveness of the solution in the context of strengthening the customer relationship can be measured in different dimensions by introducing metrics of customer and methods to evaluate these metrics. Here can be adopted metrics coming from two models RFM and RFD – recency, frequency, monetary and duration. For instance the metrics from the monetary dimension can be implemented e.g. customer value, future customer value, segment value and monitoring.

5. SUMMARY

In the paper we presented a business and technology framework for a solution for interactive marketing platform with a special recognition of acquisition, generation and management of the customer knowledge. It discusses the role of acquired, derived by business rules or calculated by analytical models data in an advanced, modern event-based and/or real time marketing

In the proposed technology platforms we introduced a lists required enabling functionality of various components which can create a complete solution for customer knowledge driven customer relationship management.

To complete analytical solution we stressed a need for strong mathematical models and other reporting tools to fulfill continuously increasing demand for adequate data both from small and big data ecospheres. It should help a process of customer relationship knowledge and to provide proper, treatment-oriented segmentation with effective and efficient propensity models.

Finally, a practical implementation of interactive marketing shows the importance of proper definition of:

- client analytical record
- propensity models with adequate granularity
- technology solution architecture
- data governance and data quality

- roles and governance process between marketing (product factory), sales distribution channels and CRM (campaign management)

REFERENCES

[1] BLATTBERG R. C., KIM B-D. NESLIN S.A, *Database Marketing: Analyzing and Managing Customers*, Springer 2008

[2] BOLJANOWIĆ J. A., STANKOVIĆ J., *The role of knowledge management in building client relationship*, Singindunum Journal 2012: 16-28

[3] CHUDZIAK J. A., CIEMSKI A., *Implementation of innovative methods of customer relationship management for marketing strategies using analytical tools, accepted for publication, 1st* National Scientific Conference, Modern problems of management, organization and production engineering, Cieszyn, Poland, 30 May, 2014

[4] FOREMAN J. W., *DATA SMART - using data science to transform information into insight*, John Wiley & Sons 2014

[5] HUGES A. M., *Strategic Database Marketing 4e: The Masterplan for Starting and Managing a Profitable, Customer-Based Marketing Program*, Mc-Grow Hill 2012

[6] MODRZEJEWSKI, J. A. CHUDZIAK, R. W. CEGIELSKI, *Complex Marketing Database Specification, Design and Implementation, Computer Information Systems and Industrial Management Applications*, 2008. CISIM '08. 7th

[7] SZCZERBA M, CIEMSKI A., *Credit Risk Handling in Telecommunication Sector*, Advances in Data Mining, Applications and Theoretical Aspects, 9th Industrial Conference, ICDM 2009, Leipzig, Germany, July, 2009. Proceedings. Lecture Notes in Computer Science Springer 2009

[8] TIWANA A., *The essential guide to Knowledge Management E-business and CRM Applications*, Prentice Hall, 2000

[9] WILDE S., *Customer Knowledge Management Improving Customer Relationship through Knowledge Applicationc*, Berlin, Springer-Verlag, 2011

Civil aviation management, soft modelling,
airline business, management concepts

Aleksander ORŁOWSKI [*]
Edward SZCZERBICKI[**]

THE CASE OF MANAGEMENT CONCEPTS IN CIVIL AVIATION BUSINESS.

The article presents the changes in management concepts used in civil aviation business in last 25 years. There is a need to make prognosis and support decision-making processes with different models but each model is adapted to current management concept. This article will present the changes and possible future trends in management concepts with the description of strengths and weaknesses of each of them. These might be used in the future works on building the models to support decision making processes which is the final goal of the future works.

The article consists of four main parts, first there is a market presented with the history and current changes. Next the main changes in management concepts are presented, basing on the example of one airline, LOT Polish Airlines. Part three consists of presentation of strengths and weaknesses of each management concept. The article is concluded and future works are presented in fourth section.

1. INTRODUCTION

This article refers to the papers presented by the authors on the previous conferences. Previously the area of interest was concentrated on the process of creation of models dedicated to help in decision making processes in the area of managing airlines and airports. Before creating models by itself it seems to be important to present different concepts of management which are in use in described business area because for each concept different type of models will be created.

[*] Gdańsk University of Technology, Faculty of Management and Economics
[**] University of Newcastle, Australia

This paper will present different concepts of management which were and are in use by international airlines. Each concept is going to be presented with strengths and weaknesses of it to get the analyze of current market concepts which are used by international airlines.

1.1. THE AIRLINE BUSINESS MARKET

The world's first airline is *Deutsche Luftschiffahrts-Aktiengesellschaft*, founded in 1909. The oldest current existing airlines are KLM and Quantas, both founded in 1920. LOT Polish Airlines was established in 1929 is one of the oldest, still operating, airlines.

In the year 2011 there were 2,8 billion passengers handled by air and more than 150 airlines which made 5,2 trillion kilometers [5].

Till the second world war, airline business was growing but still was a small part of market, a very expensive one. In the 30s of XX century it was growing so fast that for example LOT Polish Airlines was planning to open its first transatlantic (TATL) route in 1940, which did not happen due to the outbreak of World War II.

In mid 70s, British Airways started offering inexpensive transatlantic flights, which make flying popular and closed the era of transatlantic ships [6]. In the same time the deregulation of airline business has started, which led to the opening of new airlines, dropping fares and growing number of customers.

The recession in the beginning of 90s was the time in which some major airlines collapsed (PanAm, TWA). In next ten years low-cost airlines started growing in Europe changing completely the existing market. Today international airliner business is called "complex, dynamic and subject to rapid change and innovation" [1].

"A period of uncertainty" [7] is how the beginning of the twenty first century in the airline business was described by Rigas Doganis, former CEO of several airlines, currently visiting professor of the universities in the United Kingdom. Several aspects contribute to the uncertainty, first and currently most crucial are fuel prices.[8] Trying to adapt to changing market conditions airlines are modifying their management concepts. Next section illustrates changes which appear most often.

2. MANAGEMENT CONCPETS

When the authors tend to build the tool that will help supporting management processes in the company (airlines) the question starts appearing on how many aspects of business the company has direct influence. More and more services are being outsourced which should be seen from two perspectives: reducing costs versus quality control. It is also a question how can an airline try to distinguish itself when the air-

lines are outsourcing services to the same companies. That aspects will be presented in this section.

Outsourcing is the management concept that become to be used in practice in the mid 80's of XX century in the United States. The concept states "long-term, results oriented relationship of company with an external provider for activities that would otherwise provided in-house." [1]. Today outsourcing is very popular worldwide but the definition of the concept has changed in the last 25 years. There were certain steps that were done and how the concept was changing. From the theoretical perspective that types might be described:

- First: the minority of the processes were sent out of the company
- Second: also the key aspects of outsourcing were sent out of the company
- Offshoring – the process of outsourcing into the foreign countries

The section below will present the steps of using outsourcing as a management concept in civil aviation business. Because the process of outsourcing was going step by step trying to outsource more and more services it seems to be important to present the information basing on the practical examples. The authors decided to present the process of using outsourcing in the aviation business basing on the example of LOT Polish Airlines. It was done so because of two key reasons: first similar processes were observed in most western European airlines (Lufthansa, Iberia, SAS). It was done also because in the presented time period (25 years) LOT was restructurated several times and a key part of each reorganization was connected with adding new/changing current management concept used in the company. Due to that the company used most of available on the market management concepts and is still looking for the new that might be implemented.

In the beginning of the airlines business all processes connected with the air travel were owned by the airlines (starting from customer services (check in, process of selling tickets, flight attendance) to technical issues with airplanes). The process of outsourcing started slowly. First at the airports which had low frequency of flights (e.g. once a day) the airline started outsourcing the process of check-in and luggage delivery. Not to cover the cost of having the own employers in every airport which the airlines was serving the company decided to use the other companies. In the first step there were hired companies connected with the airlines being members of the same alliance (in LOT case it means airlines from Star Alliance). That means that the check-in process for LOT in Hamburg was done by Lufthansa and in Geneva by Swiss staff. In the next step LOT was using any company that was offering the check-in process and luggage delivery – that type of companies are called handling companies.

The process of replacing own employers by handling companies started moving into other airports outside of Poland. Finally LOT decided to outsource handling processes in every airport the airlines was operating, also the airports in Poland. The airline created a dedicated company (LOT Airport Services later transformed into

LOT Services) which was doing handling processes on polish airports for LOT and later for the other airlines.

As it was mentioned outsourcing handling was one of the first steps in LOT. In the same time LOT decided to outsource the whole catering process, previously the airline was producing food by themselves. There was a dedicated company established (LOT Catering) which was producing food to be served on board of the airplanes according to the menu created by LOT (and other airlines). Today the company is completely independent, it is owned by Do&Co, leading worldwide catering company and sell their products to nearly all airlines operating airports in Poland.

Similar processes happened with offices in which tickets were sold. It used to be done in the way that an airline had an office in every city that the airline was serving. In the middle 90's of twenty century offices started being closed. Tickets were sold through associates travel agencies, using call centers and via Internet (here also mostly via associated travel agencies). Today LOT closed almost all offices, still having them just in few cities in Poland.

Outsourcing was not only connected with products and services which are not building the "key value" for the customers (as the original idea of outsourcing says). In December 1996 LOT Polish Airlines established a new company: EuroLOT. It was a wholly owned subsidiary of LOT Polish Airlines. The main goal of the company was to operate on the domestic and short-houl routes (up to 1,5 hour flight) for LOT Polish Airlines. Having the new company LOT was reducing the operational costs of flights. But also in that situation part of the company's core business (flying) was sent out of the company. In the year 2011 LOT sold Eurolot. Next LOT decided to let the company totally operate all domestic flights, buying just certain number of seats in each flight (all done to reduce the costs). Finally these changes were not implemented because after doing so LOT would have no control on their network. Having a main airport, hub, in Warsaw, LOT is interested in transferring passengers from regional airports via Warsaw to other places in the world. In that strategy it is important to coordinate incoming flight and next flight flying out of the airport. In presented case Eurolot was suppose to build their own schedule of flights from regional airports to Warsaw. Practically it meant that flights might not be connected e.g. the passenger was flying from Gdansk via Warsaw to Milan Malpensa. The flight from Gdansk to Warsaw arrives Warsaw at 9 a.m., passenger has to wait in Warsaw 5 hours (because the flight to Milan Malpensa left 30 minutes earlier) and then spend 2 hours on the way to Milan Malpensa. Today Eurolot as an independent airline operates flight both for LOT in ACMI (the lessor provides aircraft, crew, maintenance and insurance)/ Wet Lease (used interchangeable with ACMI) formula [4]. Eurolot is an example of regional airline so as City Hopper (operating for KLM) or Wideroe.

One of the key aspects of each airlines used to be aircraft maintenance. 2010 LOT established LOT Aircraft Maintenance Services (LOT AMS) and moved all technical employees to that company. Starting from that time all airplane services are done by

companies which are not owned by LOT, today not even by LOT AMS (new airplanes, Boeing 787, are under technical supervision of Monarch, one of the leading worldwide companies, not LOT AMS and all technical issues are done out of Poland).

Starting from that point the question appears "is it possible to outsource anything more?". LOT, as most of the airlines these days, does not own the airplanes, all airplanes are leased. It might be stated that this situation has influence only on the financial aspects of the company and have no negative influence on customers relations neither the quality of services offered by the company. The airplanes are leased from other companies. LOT also used "sales-and-lease back' concept in two areas of business: with airplanes which used to be owned by the airline but also the headquarter of the airlines was leased in that way. It seems to have just the financial aspect helping the airline gain the capital.

There are market examples in which airlines are trying to outsource the crew, mostly using offshoring (outsourcing the services abroad). Norwegian Air Shuttle [2] registered a company in Ireland where labor regulations allow foreign flight attendants to work for less. Due to that the airline hired flight attendants from Thailand who are working on board of airplanes operating from Oslo to the United Stated and are paid $500 a month. Similar strategy is trying to be adapted by Finnair, the airline plans to outsource cabin crew on 43 routes and due to that reduce costs by 18 million Euro/year. Other European airlines (e.g LOT) are analyzing that concept.

Looking at the examples presented above the company owns only the brand, management systems and management system. All other aspects are outsourced to other companies. Trying to find the answer which type of structure in airlines might be expected in the coming years it seems to be necessary to present the example of the Middle East airlines.

2.1. OTHER CONCEPTS

Emirates is one of the fastest growing airlines in the world, having 218 airplanes and serving 134 destinations and more than 44 million passengers worldwide (June 2014) [3]. The airlines concentrates on the quality of the product, due to that the company owns 32 own airport lunges located in 28 cities in the world. The airline keeps their own staff for first and business class passengers at the airports the airline is operating to. Also the staff is fully hired by the airline. But still it should be mentioned that Emirates Group has more than 50 subsidiaries (e.g. catering company, engineering company) so the services are outsourced but mostly to the companies which are under control of Emirates Group. This is because Emirates tries to deliver superior customer service as a key strategic point and it is easier to control the quality of services in one group of companies.

While observing the European civil aviation market there is also one example of different management concept. Czech Airlines (CSA) were merged with the company

administrating the main airport in Czech Republic, Prague Ruzyne. It might look like the opposite process to the steps of outsourcing presented before. It must be mentioned that this process was done due to very bad financial results of CSA and very stable situation of the airport. Practically it might be concluded that the airport pays for the airlines debt. This process was reported to the European Commission by other airlines operating that airport and should not be presented as the management decision but clearly political decision and due to that it is not going to be analyze in the next sections.

3. CONCEPTS EVALUATION

In the previous section there were presented different steps in the concept of outsourcing which are used by international airlines. It seems to be necessary not only to present the concepts but also try to evaluate them presenting strengths and weaknesses of each. Having that information it will be possible to plan the usage of models supporting management in each of the presented concepts.

3.1. OUTSOURCING SIMPLE PROCESSES

First the processes of outsourcing basic, not creating key-value, processes is going to be evaluated. From the processes described in the previous sections outsourcing of handling, catering, ticket offices and sales and lease back are going to be presented.

Outsourcing catering and handling will be presented together. Both products might be part of the service that is sold by the airline (luggage, food on board of the airplane), customers review the airline by the quality of that products but it is easy to find/change the outsourcing partners (because of large number of catering/handling companies)

Table 1. Strengths and weaknesses of outsourcing of handling and catering

Outsourcing handling, catering	
Strengths:	Weaknesses:
Lowering costs (the percentage depends on how many flight were operated to an airport on which home handling was existing, the lower the frequency of flights is the higher savings might be gained.	Customers do not know that the airlines is not responsible for lost luggage or for the check-in process. So all negative aspects of handling will influence on the airline but the airlines does not have full influence on handling.
	Less control on the processes by airlines

Taking into consideration presented weaknesses the airlines prefer to outsource handling and catering because of high financial savings.

The processes of outsourcing ticket offices seems to be more management/logistics aspects of the airline business then catering/handling. It is changing the way of thinking about the service (selling tickets) not the service by itself (like in catering/handling).

Table 2. Strengths and weaknesses of outsourcing ticket offices

Outsourcing ticket offices	
Strengths:	Weaknesses:
Lowering costs (selling the ticket on airline web page can be up to 100% cheaper than in the city office)	Limited contact with client – if the airline does not have their own selling ways (no difference if these are city offices or company's webpage) there is limited flow of information
Possibility to use bigger number of selling points (travel agencies)	Selling tickets through other companies is expensive (commission must be paid) when compared with direct distribution via airlines web page
	Outsourcing partners normally sell tickets offered by different airlines, not that much concentrated on the tickets of one airline

Outsourcing ticket offices can limit the costs that strongly that airlines mostly decided to use that management concept. Having in mind weaknesses special programs for travel agents to attract the airlines tickets are created.

Table 3. Strengths and weaknesses of sales and lease back

Sales and lease back	
Strengths:	Weaknesses:
It is possible to use the capital which was invested in airplanes or buildings	In total more expensive then owning the building or airplane
Normally cheaper than bank loan	
If the company's financial condition is bad that is the only source of capital that might be used	

So as in previously presented outsourcing concepts, the financial advantages are so high, with limited weaknesses, that most airlines looking for a capital decided for that concept.

3.2. OUTSOURCING KEY AREAS OF BUSINESS

Outsourcing key areas of business is the second step in implementing outsourcing as a concept of management. In the area of civil aviation it mostly mean establishing new airline and that example was described in the previous section. Here the strengths and weaknesses of that process are going to be presented.

There might be two main strengths of the idea presented, first is lowering the operating costs of an airline. Due to the fact of creating new airline, mostly operating smaller, regional airplanes there are none union agreements (as in original airlines) and there is very limited administration. This argument is connected with the second advantage of establishing new airline. Using that idea is easier to control the total costs of the flights or group of flights. Normally there are certain variables influencing the total cost of the flight: crew, fuel, taxes, maintenance, flyiing rights, cost of landing, handling, airplane (leasing) which makes the calculation of the single flight costs complicated. Outsourcing the flights to a new company makes the process easier because there is a need to buy a service "flight" not thinking about sub costs mentioned above which has to be calculated by established company.

Establishing new company may have several weaknesses. If the newly established airlines is controlled by main outsourcing partner the possible problems presented below will appear. First if the airlines are totally independent there is a problem with coordination of flights (which is very important for legacy airlines operating hub and spoke model). This problem is fundamental for the possible success of an airlines. It should be mentioned that there might be quality standards problems while independent airlines might follow different quality standards. That also means different passengers regulations like luggage allowance, catering, loyalty programs.

The process of full outsourcing part of core business is very risky with strong influence on the most important managerial decisions. Because of that very few airlines decided to do so, mostly the airlines with huge financial problems for which short term financial arguments were most important.

3.3. FUTURE PROCESSES

As it was presented in the previous section there are some management concepts which starts appearing on the market. The most important seems to be the process of outsourcing cabin crew.

Airlines decide to outsource the crew because of two most important aspects. First is the possibility to lower personnel costs, due to that first airlines that started that process are located in Northern Europe (Norway, Finland) where the personnel costs are the highest in Europe. Second biggest advantage is the flexibility of staff, airline business is strongly connected with economic circles. If there is a need to reduce the number of employees it is easier (faster and cheaper) to do so with employees who are

outsourced than with the people hired directly by our organization. That two arguments are so important that will probably attract that management concept to other airlines

Outsourcings crew has some important disadvantages. First there are several law problems with hiring personnel in that way (e.g. conflict between Norwegian and US Transportation Department). Second is the problem with setting company's standards, especially quality standards which will be more complicated to implement and control when the employees are not directly employees of an airline. Finally the serious managerial problem appears, if handling, ticket offices, call centers and crew will be outsourced that might mean that the customer will meet none of the airline employees in the whole processes of contact with an airline. That will bring the problem how to distinguish from all other airlines.

3.4. OVERVIEW

Each of the presented concepts might be seen from the perspective of its strengths and weaknesses. Having that it is possible to build the analyze (even so basic as SWOT) but each analyze will have to strongly base and be seen from the perspective of the airline for which it is going to be created. There will be strong differences between the airlines (e.g. between SAS and Lufthansa due to huge difference in personnel costs) so it cannot be presented as a general statement for all airlines. Due to that the authors decided just to point strengths and weaknesses of each concept which might be used in analyzes created for each airline.

Because the market has been observed by the authors for a long time it seems not to be the point how many things might be outsourced, basing on the presented examples (Norwegian) almost all aspects of business might be outsourced. It does not mean that all airlines will follow that management concept. The airlines will try to chose the concept which will bring high flexibility with the possibility to keep quality level and building competitive edge. Currently nearly all airlines in the world outsource simple processes, most also established the new airlines to operate regional destinations. Nothing more in general can be stated today, most of regional airlines are still owned by the airlines which established them mostly because of weaknesses mentioned in that section. According to official reports the most important airlines (Lufthansa, AF/KLM) are not going to change that concept. But airlines are very interested in outsourcing crew because of both financial savings and flexibility and it seems as this concept might get popular in the coming years.

4. CONSLUSIONS AND FUTURE WORKS

The article presents the way how management concepts are used and have changed civil aviation business in last 25 years. The goal of the paper was to define and describe management concepts which are currently in use or might be used in the mid-term. The type of concept might strongly influence on models which will be described to support decision making processes on that market.

The first part of the paper presents the description of the market with its history. Next the management concepts which appeared on the market are presented. The concepts are presented in chronological order, basing on the example of one airline: LOT Polish airlines. Next section presents the assessment of each of the described concepts presenting strengths and weaknesses of each.

There is no answer given which concept is the ideal one because it strongly depends on the airline, its position, costs and strategy. But the presentation of strengths and weaknesses should help the airlines managers to make the decision which concept the airline is going to adapt. When that decision is done it will be possible to create a model that will support decision making processes for the created organization structure.

Future works will concentrate on creating one model supporting management decisions (predicting number of passengers on newly created destinations) for dedicated airline. It will take into consideration the management concept (and as a result organization structure) planned for the airline in mid-term perspective.

REFERENCES

[1] DOGANIS R., *Flying Off Course, Airline Economics and Managent,*Fourth Edition, London and New York, Routledge 2010, 1.

[2] http://www.rollcall.com/news/us_carriers_wary_of_norwegian_airlines_cut_rate_wages-229935-1.html 30.06.2014

[3] http://www.emirates.com , 30.06.2014

[4] http://www.globalplanesearch.com/aircraft_leasing/definition.htm, 30.06.2014

[5] IATA Annual Report 2012, http://www.iata.org/about/Documents/annual-review-2012.pdf, 1.7.2014.

[6] http://traveltips.usatoday.com/history-airline-industry-100074.html , 1.07.2014.

[7] DOGANIS R., *The Airline Business,* Second Edition, London and New York, Routledge, 2005, 258.

[8] ORŁOWSKI A., SZCZERBICKI E., *Modelling processes for the Polish civil aviation market,* Knowledge Based Approach to the Design, Control and Decision Support, 2013, 33–42.

Model, accounting system, standard software

Paweł KUŻDOWICZ*

MODELING THE ACCOUNTING SYSTEMS IN STANDARD SOFTWARE

The article presents selected aspects of modeling accounting systems in the enterprise. It focuses on modeling data flows in the ERP (Enterprise Resource Planning) system in the context of recording costs and achievements. Selected flows in the area of sales, purchasing and production have been characterized taking into consideration the accounting records on general ledger accounts and on controlling objects. A concept of presenting and reporting flows has been demonstrated in a table format, linking the general ledger (GL) with controlling objects.

1. INTRODUCTION

The current changes in the accounting development trends result from changing information demand from stakeholder groups, on one hand, and the dynamic development of IT which brings new recording and reporting capabilities, on the other. The sophisticated architecture of standard management software, especially the increasingly popular Enterprise Resource Planning systems, allows for the development of new recording models [Saniuk A., Cagáňová D., Čambál M., 2013]. These models are created to implement accounting systems that include the General Ledger (GL) and the enterprise costs and achievements account objects.

The purpose of this paper is to present the concept of modeling accounting systems embedded in an ERP system. Emphasis was placed on selected value stream flows in the sales, purchasing and production area, taking into consideration the accounting records on general ledger accounts and on costs and achievements account objects. To meet the aforesaid objective, literature study and observation research methods were used.

* University of Zielona Góra, Faculty of Economics and Management.

2. IDENTIFICATION OF VALUE STREAM FLOWS

Modeling accounting systems in an ERP system consists in identifying the value stream flows. The source of these value stream flows in an enterprise is logistics, i.e. logistic processes which are subject to proper recording and reporting in the enterprise bookkeeping system, which includes the (financial) accounting subsystem and the controlling subsystem. Accounting is responsible chiefly for preparation of the general financial report at the end of the financial year, while Controlling is responsible for internal detailed reporting also in shorter cycles (monthly and quarterly) [Kużdowicz P., Witkowski K., Vidová H., 2013].

2.1. THE NOTION OF VALUE STREAM FLOW

A value stream flow in the logistic chain includes changes of expenditures (costs) and attainments (results) occurring at the individual links of the chain in connection with execution of the procurement, production, and sales processes applied to the manufactured goods. Such changes may be identified in terms of various categories (e.g. expenditure types, attainment types) and/or analytic objects (e.g. cost centres, profit centres, products, customer groups). Changes of expenditures are typically accompanied by quantitative (material) movements, such as inventory increase or consumption of production means [Kużdowicz P., Relich M., Kużdowicz D., 2013, s. 336]. For example, when material is released for consumption in production, there is a material movement, i.e. the inventory is decreased and consumption is increased. On the other hand, changes of attainments may be connected with material movements (e.g. inventory change) or not (e.g. issuing an invoice to a customer). As the execution of production processes progresses, expenditures and attainments change at successive links of the supply chain. Releasing materials for production results in increasing the value of consumed materials and of the production in progress, at the same time decreasing the value of the material inventory. As production orders are closed, the final-product inventory increases and the value of production in progress decreases. Delivery of final products to the customer results in increasing the value of sold products and decreasing the value of final-product inventory.

2.2. IMPACT OF ACTIVITY TYPES

The development of the enterprise accounting model, especially the controlling subsystem, depends mainly on the enterprise size and activity type. These factors translate into controlling types. The most common include [Hölzlwimmer, (2009). s. 39]: customer order controlling, production controlling and procurement controlling. Manufac-

turers of industrial unit appliances (installations) usually perform customer order controlling where focus is on determining and posting costs in respect to a customer order (purchase order). In the case of projects extending over many months or even years, the use of prepayments and accruals is necessary, as it allows determining the current costs and revenues before the order completion. Production controlling is important in the case of a serial production of goods delivered to a warehouse. Calculation and settlement of general costs and the recording of warehousing costs play a major role here. In turn, commercial companies do not need production controlling. Instead, they focus on the purchase and sale of commercial goods, concentrating on supplier management and controlling, and on customer profitability analyses. Purchasing costs and sales turnover are of key importance for maximizing the profit of a commercial business. However, it does not mean that a specific organization relies on a single type of controlling. Frequently, all of them are applied at the same time to a various extent, depending on the condition and development stage of the entire organization.

2.3. INTEGRATED APPROACH

In respect to business applications of standard software, integration means an automatic combination of different organizational areas, such as production, sales or finance. Its purpose is to mutually combine data originating from various sources and deliver them to stakeholders. Identification of the value stream flows in the context of information needs of various stakeholder groups in an enterprise requires an integrated approach. Let us consider this using the example of internal stakeholder groups, among which the Logistics, Accounting, and Controlling departments are identified (cf. Hölzlwimmer 2009, p. 29). The requirements from Logistics focus on material movements. Subject to analysis is all information which ensures smooth execution of the production and delivery processes. Such information may come from the product-cost accounting provided by Controlling. On the other hand, Controlling relies on information from Logistics, e.g. for the purpose of product calculation. Such information is related to technological specifications (recipes) or product structures as documented in the production plan. The primary task of Controlling is to timely detect any bottlenecks in the enterprise in order to undertake corrective measures. The Accounting department „consumes" the data from Logistics: logistic transactions are subject to valuation for the purpose of their assignment to the general-ledger (GL) accounts and cost objects.

The process of integrating information resources and IT services within an organization is a permanent phenomenon visible through the prism of implementing the entire family of integrated systems. Therefore, each organization seeks to strengthen the potential behind the integration by combining various solutions to form a whole [cf. Zaskórski, 2012, p. 203].

3. MODELING VALUE STREAM FLOWS

To identify value stream flows in a logistic chain, mapping models are applied which enable the analyzed flows to be illustrated graphically. Among the most frequently used are Porter's value chain models and the SCOR (Supply Chain Operations Reference) model developed for analyzing supply chains and identify potential improvements in flows of goods, labor, and information. The remaining part of this paper focuses on examples of application of the COR model for identification of value stream flows in procurement, production, and sales processes. Flows in individual processes are analyzed both from the point of view of Logistics (material movements) and from the point of view of Accounting and Controlling (value stream flows). Processes are always mapped through relevant documents (receipts etc.).

3.1. VALUE STREAM FLOWS IN PROCUREMENT PROCESSES

Fig. 1 shows the possible documents generated in procurement process recording. The following logistic processes in the Procurement area are distinguished: purchase order, stock receipt, bill input, and payment output. The documents are accompanied by respective streams of primary value flows in the financial areas (Accounting and Controlling).

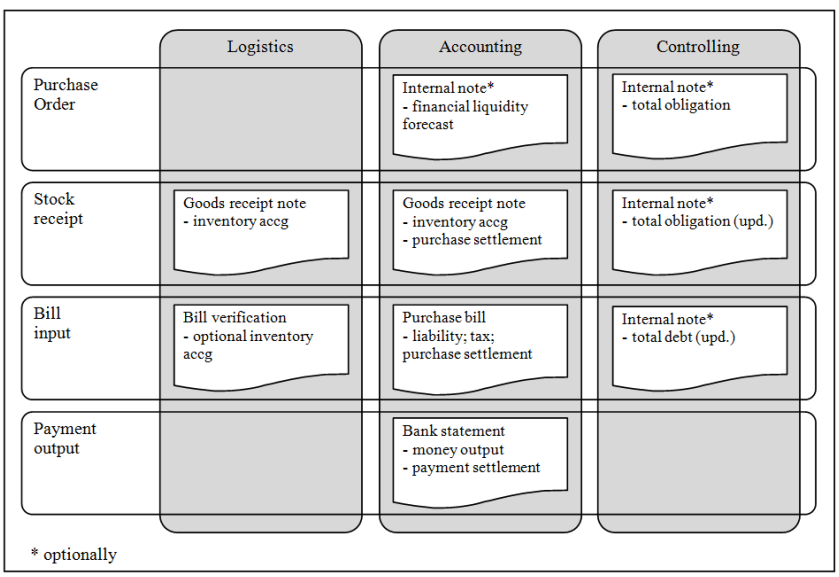

Fig. 1. Value stream flows in procurement processes [Hölzlwimmer 2009, p. 93]

The flows of the contemplated streams in the Accounting area are obligatory. They include posting the following to GL accounts: stock receipts, purchase bills, payments of the obligation. Flows in the Controlling area are optional and include in particular recording the total obligation in connection with placing orders for materials, as well as its updating in course of subsequent (follow-up) processes. Individual processes are triggered automatically by recording the source documents in the Logistics area: internal obligation note (at issuing the order), goods receipt note (at receipt of the goods), purchase bill (at approval of the controlled bill). Individual streams are posted to GL accounts and cost objects (as appropriate) through default account assignment defined at the moment of recording the order or any changes to subsequent (follow-up) documents. An exception is the bank statement document whose source is outside Logistics and which is recorded only in Accounting.

3.2. VALUE STREAM FLOWS IN PRODUCTION PROCESSES

Fig. 2 shows the possible documents generated in production process recording. The following logistic processes in the Production area are distinguished: production order, requisitions (material)data entry, production data entry, and completion notices.

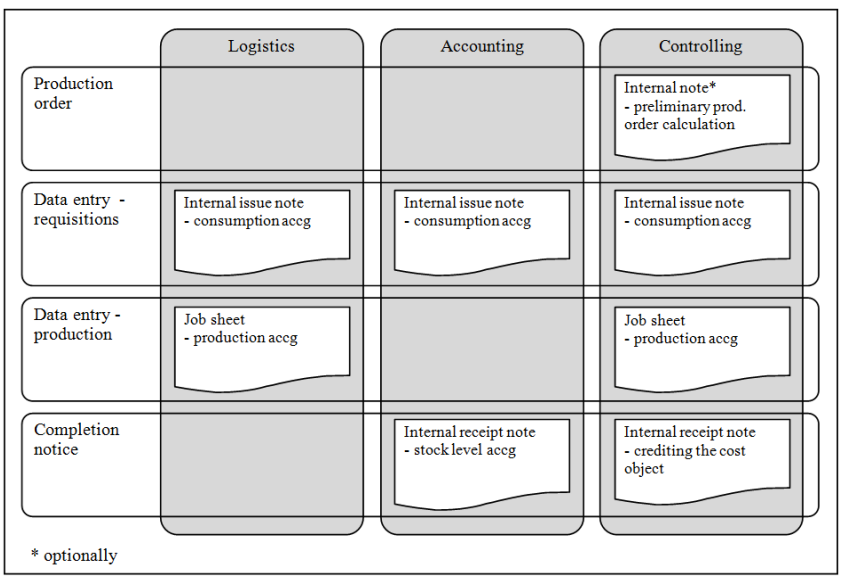

Fig. 2. Value stream flows in production processes

It is assumed that the production process is triggered by the production order document. It does not generate any flows in the Accounting area, but optionally enables per-

forming preliminary production order calculation in the Controlling area. The first document generating value stream flows on GL accounts is the data entry requisitions note. In terms of value, it is credited to the inventory account and debited to the material consumption account, at the current inventory price of the given product. At the same time, the cost object assigned to the production order is posted. Data entry production reports are based on so-called job sheets and reflect completed working operations. The accounting is performed through cost objects, using the direct-labor and machine-hour cost types. Therefore, production reports generate value stream flow only in the Controlling area, bypassing the Accounting area. The document which terminates the process contemplated here is the completion report which generates value stream flows in both Accounting and Controlling. Receipt of the final product to the warehouse (at a predefined value, e.g. current cost, standard cost, postulated cost) is posted to the GL accounts. Additionally, the cost object is debited (at the same value).

3.3. VALUE STREAM FLOWS IN SALES PROCESSES

Fig. 3 shows the possible documents generated in sales process recording. The following logistic processes in the Sales area are distinguished: customer order, external shipping, sales invoice, and payment input.

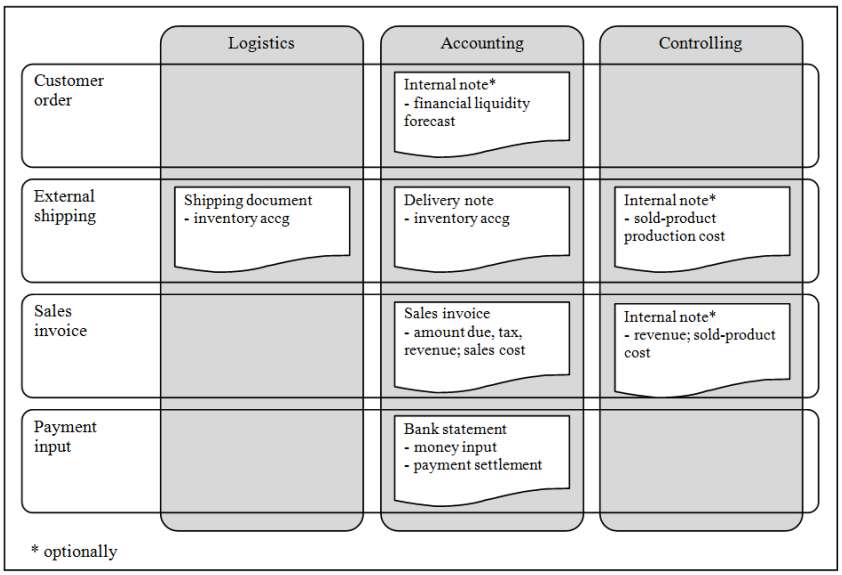

Fig. 3. Value stream flows in sales processes [Hölzlwimmer 2009, p. 172]

It is assumed that the sales process is triggered by the customer's order document. In practice, it may be preceded by a request for proposal and a proposal, which documents do not affect the contemplated flows after all. While the customer order does not generate entries in GL accounts, it is visible in the financial liquidity forecast. The first sales document generating value stream flows on GL accounts is the shipping document. The inventory account is credited and the sales settlement account is debited at the current inventory price of the given product. Optionally, the relevant cost object may be also debited. The subsequent (follow-up) document is an invoice which triggers postings to GL accounts (amount due, tax, revenue, product sales cost) and to cost objects (e.g. revenue, production costs of sold products by customer, region, or segment etc.). Also, the invoice closes the shipping document by generating additional entries on the sales settlement accounts in the GL. The final document in the contemplated process is the bank statement which confirms that the payment has been effected and settled. In case of discount, additional accounting entries in Controlling are optionally generated, decreasing the postulated revenues (Kużdowicz P., Witkowski K., Vidová H., 2013).

4. COSTS AND ACHIEVEMENTS REPORTING

From the value stream flows perspective, accounting costs and achievements in an ERP system consists in recording business operations on general ledger accounts and costs objects at the same time. The concept of reporting those flows in the table format, compared to the classic format, is presented later in this paper.

4.1. TABLE FORMAT

In ERP systems, financial *data* are presented and reconciled in the table format with rows corresponding to specific cost and revenue types, and columns corresponding to cost centres and drivers and results (cf. Figure 4). The cost centre (CC) collects costs by place of origin (e.g. procurement or production department). The cost object (CO) collects unit costs of an item which is the reason behind the cost (e.g. a product, customer order, warehouse order or project). The income object (IO) collects revenues and costs of the period concerned which generate the result of business operations (e.g. customer, item group or sales region).

Usually, the illustration of specific items starts with income/revenue rows, followed by rows corresponding to various cost items. The full approach is to separate the result on core business operations, and cost-neutral / revenue-neutral settlements and by-function settlements. Later in this article, we will not provide an in-depth analysis of those settlements. Instead, we will focus on the area of core business costs and achieve-

ments accounting. Depending on their type, items may be allocated to 2–4 areas at the same time. For example, the consumption of primary materials as a direct cost (in the amount corresponding to the volume of recorded internal consumption) in general ledger accounts and cost driver accounting objects at the same time; depreciation as expenditure / general cost in general ledger accounts (in the amount of balance sheet depreciation) and in the cost centre account (in the amount of functional depreciation) at the same time.

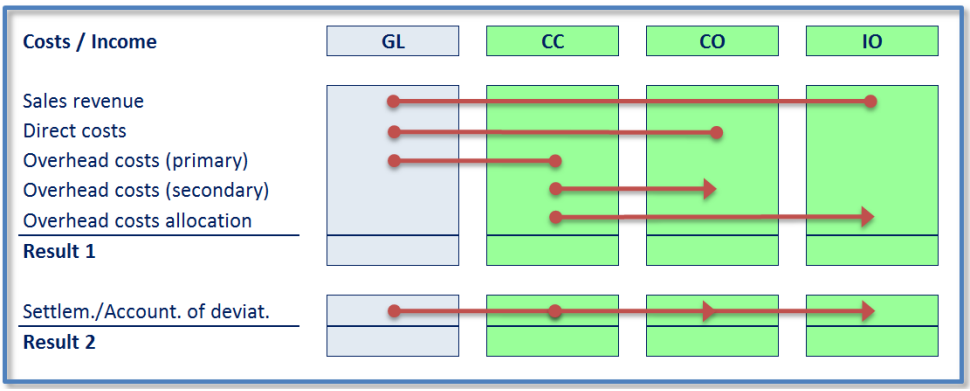

Fig. 4. Identification of value stream flows in the table format

In the table format, the **reporting period closing procedure** means determining subsequent results cumulatively. **Result 1** is the difference between sales revenue and primary materials costs, primary production costs (e.g. labor costs and machine-hour costs) and general costs incurred. Sales revenue, primary materials costs, and primary production costs and other general costs are allocated directly to income objects, cost drivers and costs centres, respectively. The result also takes account of cost centre crediting. This includes secondary production costs and general costs (e.g. overhead on purchasing, production, and sales costs), and is generated to additionally debit the cost drivers. The last step is to calculate other costs which simply involves clearing the cost centres to zero and adding the relevant amount to cost drivers or results, which helps determine **result 2.** Clearing the cost centres to zero is related to posting deviations in the general ledger accounts. In practice, this depends on the available software [Kużdowicz, Kużdowicz, 2013].

4.2. EXAMPLE

Below, you may see an example showing the advantages of the proposed approach. It presents the method for identifying selected value stream flows both in the classic and table format. The following business operations, which represent the primary value

stream flows, are considered: costs by type in the following amounts, respectively: €60 k (*1a*) for raw materials consumption; €50 k for primary production costs (*1b*) and €20 k for other general costs (*1c*); receipt of internally manufactured products: €100 k (*2*); release of sold products: €80 k (*3*) and revenue from sale of products: €150 k (*4*). Secondary value stream flows are applicable to operations involved in closing the reporting period. In addition, operations related to the control over the chain of cost accounts (9a–e) and to the transfer of sales costs and revenue to the financial result account (9f) were identified. A detailed presentation of other assets and liabilities concerning receivables and liabilities etc. was omitted in the example.

CLASSIC FORMAT

In the classic approach, the contemplated value stream flows are recorded in Group 4 and Group 5 GL accounts (this includes accounting by type and/or by function, cf. Figure 5). The settlement within the closing of a reporting period means using a single chain of cost accounts.

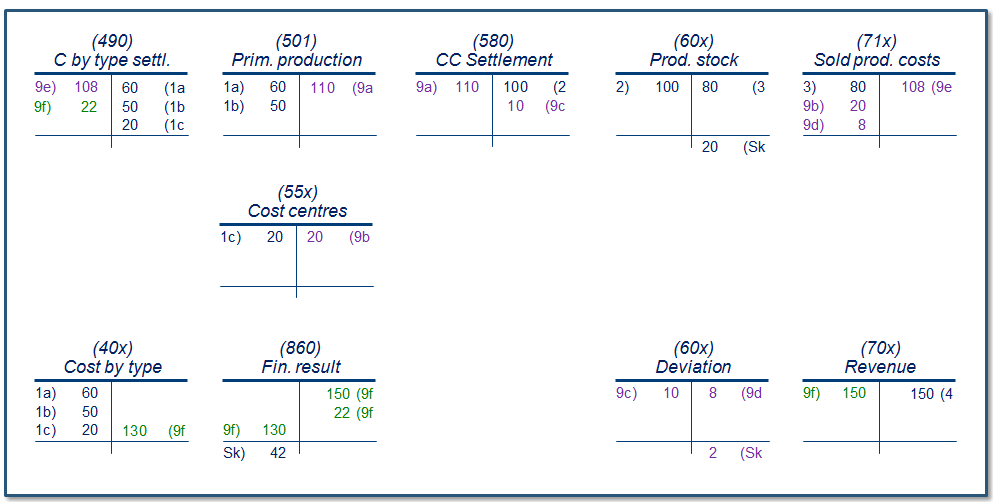

Fig. 5. Presentation of value stream flows in the classic format

In the general ledger, within the closing of a reporting period (operations 9a–f), costs and sales revenue were recorded in amounts of €130 k and €150 k, respectively. The recognition of change in the balance of products in the amount of €20 k (adjusted for deviation of €2 k) allowed determining the sales result in the amount €42 k in the reporting period concerned.

TABLE FORMAT

The table format, used in standard ERP software, includes recording business operations both in GL and in the costs and achievements account as part of the double chain of cost accounts. See Figure 6 for the results of flows considered in this paper.

Costs / Income (in k€)	GL	CC	CO	IO
Sales revenue	-150			-150
Increase product stock	-100		-100	
Sold product costs	80			80
Raw materials	60		60	
Labour expenses/costs	50	50		
Other overhead costs	20	20		
Manufacturing costs (secondary)	0	-40	40	
Overhead costs allocation	0	-20		20
Result 1	**-40**	**10**	**0**	**-50**
Settlem. of prod. costs deviat.	0	-10	2	8
Account. of prod. costs deviat.	-2		-2	
Result 2	**-42**	**0**	**0**	**-42**

Fig. 6. Presentation of value stream flows in the table format

Due to single-sided data presentation, achievement amounts (income/revenue/result) are prefixed with a "-." **Result 1**, in the total amount of €-40 k, is calculated by recording sales revenue in the amount of €150 k, the change in the balance of products in the amount of €-20 k (= -100 + 80), primary materials costs in the amount of €60 k, primary production costs in the amount of €50 k, and other general costs in the amount of €20 k. Result 2 takes account of crediting the cost centres and cost drivers, and debiting the result drivers. Once credited with secondary production costs in the amount of €-40 k and with other general costs in the amount of €-20 k, the cost centres provide a result of €10 k. The results on cost drivers and result drivers are €0 k and €-50 k, respectively. Recording the abovementioned operations caused a change to the result's allocation to cost centres, cost drivers and result drivers. Due to their achievements, the cost centres were credited, and the result was presented as the deviation amount. Creating and receiving the products in the warehouse resulted in clearing the cost drivers to zero. This means the outstanding deviations include only production costs (including, without limitation, labor costs). This allocation of costs allowed determining the result for result drivers which only needs to be settled due to existing deviations. **Result 2**, in the total amount of €-42 k, was obtained by recording the deviation in the amount of €10 k identified in cost centres. This amount was settled by crediting the costs centres and debiting

cost drivers and result drivers (pro rata to sold product costs) with the amounts of €2 k and €8 k, respectively.

This example shows the method for identifying value stream flows corresponding to basic business processes implemented in the enterprise. Flows generated between specific GL objects and costs and achievements objects ensure permanent (continuous) compatibility between general ledger and cost account areas at various analytic levels. The analytic levels identified in this example, corresponding to particular results, specifically allow assessing the efficiency of cost centres (result 1) and sales profitability of various result driver groups (result 2). The use of the table format also ensures adequacy between planning data (which is actually summarized in tables) and execution data.

CONCLUSIONS

The discussed approach shows an alternative recording and reporting method in an enterprise accounting system embedded in standard software. This means recording value stream flows that link the general ledger with costs and achievements objects (i.e. cost centres, cost drivers and result drivers). The table format implemented in ERP systems was used. While addressing the needs of various stakeholder groups, it ensures adequacy between planning data (summarized in tables) and execution data. The usefulness of the contemplated approach depends on the way the flows are modeled for the purpose of their later analysis, and may vary in function of the organization's growth and market evolution. This means the need to adopt a situational awareness approach, i.e. the need to adapt the deployed solution on a continuous basis.

Further studies will be aimed at developing a method for the modeling of value stream flows in the context of their continuous adaptability to the evolving environment and to changes in the organization itself. Therefore, the proposed solution should keep up with the changing and improving supply chain of the organization.

REFERENCES

[1] HÖLZLWIMMER A., (2009), *Integrierte Werteflüsse mit SAP ERP*, Galileo Press, Bonn.

[2] KUŻDOWICZ P., RELICH M., KUŻDOWICZ D., *Przepływy strumieni wartości w łańcuchu dostaw przedsiębiorstwa*, Logistyka 5/2013, ILIM Poznań, 335-340.

[3] KUŻDOWICZ P., WITKOWSKI K., VIDOVÁ H., *Modelling value stream flows in the enterprise supply chain*, [w:] Management 2013, Vol.17, No. 2, , pp. 205-217, ISSN 1429-9321.

[4] SANIUK A., CAGÁŇOVÁ D., ČAMBÁL M., *An IT solution for supporting production order planning in make-to-order systems of SME's*, in: Logistyka, 2013, nr 5, 373-379.

[5] ZASKÓRSKI P., *Integracja zasobów i usług informacyjnych w organizacji biznesowej*, Zeszyty Naukowe WWSI, nr 7/2012, Warszawa 2012, 203-217.

Assessment financial condition, standard assess, SME

Zofia WILIMOWSKA[*], Agnieszka PARKITNA[**]
Arkadiusz GÓRSKI[***], Anna Maria KAMIŃSKA[****]
Anna KIŁYK[*****], Katarzyna GWÓŹDŹ[******]

ASSESSMENT OF THE FINANCIAL CONDITION OF SMES: IS THERE A NEED TO ESTABLISH NEW COMPULSORY OR VOLUNTARY FINANCIAL STANDARDS?

This paper contains a discussion about the problems of the financial condition assessment standards in the SME sector in Poland. This work attempts to analyze and diagnose the state of the existing literature and business practice. The article suggests the need to adapt to the demand of the practice of literary activity, by identifying shortcomings. The result is to create a plan of research methodology, which would solve the current problems.

1. INTRODUCTION

The issue of assessing the financial condition of SMEs is an important subject because of its role in scientific development of finance, as well as the importance of this sector. Analysis of the financial condition of the SME sector at present weakly researched, and it is necessary to conduct actual research to develop adequate tools, methods, and standards used in the evaluation of companies in this sector. The SME sector have a huge impact on economic development and innovation of any economy, also in Poland. According to the European Commission in 2014 in the SME sector will be about 6 million workplaces [14], which is three times more than in the large enterprises. Sector of small and medium enterprises is a sector, where companies do

[*] Wroclaw University of Technology, e-mail: Zofia.Wilimowska@pwr.edu.pl
[**] Wroclaw University of Technology, e-mail: Agnieszka.Parkitna@pwr.edu.pl
[***] Wroclaw University of Technology, e-mail: Arkadiusz.Gorski@pwr.edu.pl
[****] Wroclaw University of Technology, e-mail: Anna.Maria.Kaminska@pwr.edu.pl
[*****] Wroclaw University of Technology, e-mail: Anna.Kilyk@pwr.edu.pl
[******] Wroclaw University of Technology, e-mail: Katarzyna.Gwozdz@pwr.edu.pl

not prepare synthetic reports for financial institutions such as: the balance sheet, profit and loss account or cash flow.

It is almost impossible to calculate the commonly used financial ratios and requires solutions for assessing the financial standing of the SME sector - need to take into consideration this restriction. An effect of flexible economy is a large variety of business profiles of SME sector companies. Flexible economy has results in a large variety of business profiles SME sector companies. This therefore requires different approach to the different groups of firms, and hence the use of specific standards of analysis on which the vote would be tested by the operator.

2. THE ROLE OF SME SECTOR

Number of all active enterprises gives Poland sixth place among the countries belonging to the European Union. In Poland, as in the European Union, SMEs are approx. 99% of all active enterprises. The share of large firms is therefore negligible. The difference between Poland and the European Union refers to micro-enterprises, which number in the country exceeds the European average (96% of Polish companies are micro-enterprises, in the EU the figure is 91.8%).

During analyzing the structure of industry of domestic enterprises it can be concluded that they operate commercial, service and activities related to construction. Less common are large companies. In 2012, as in previous years, most companies rose in the area of trade and construction (respectively 30.1% and 13%). Until 2008, the second most popular section of PKD for start-up businesses was an area of finance and insurance.

According to Eurostat, in 2008 68.9% of working Poles were employed in the SME sector. The share of people working for micro but was 39%. Less than a third of the 8.8 million Polish workers have worked for large corporations. Data from the Central Statistical Office (CSO) of the same year reported that every second person has found a job in the micro and small enterprise (52%) and a fifth (19% of respondents) in the medium company [2]. Data from these two sources are therefore comparable. In terms of the structure of employment situation in Poland is comparable to the situations in the countries of the European Union. However, there is a significant difference in the size of the SME sector. In terms of number of employees per company Poland ranks only 18th place. In the country on a single economic entity accounts for 4.4 people (for the SME sector, this result stands at 3, 8 employees per company) [11].

In 2011 we conducted a study on the economic activities of companies employing up to 9 people. In that period led 1772.6 thousand. Entrepreneurs - there was a 3.3% increase in the number of entities compared to 2010: 94.4% (1672.8 thousand.) units of this nature are natural persons. The remaining share belongs to legal persons and entities without legal personality (99.8 thousand. Micro). There were micro-employed

3578.1 thousand, of the average monthly gross wage per employee was 1987.4 PLN. Most businesses were related to commercial activity and repair of motor vehicles and construction (respectively 33.2 and 13.1%) [1]. In these areas, micro, were granted the highest revenues. The rest of the activities of micro-generated revenue of 816.5 billion PLN, while expenses amounted to 716.7 billion PLN level. Businesses have invested in their activities until 25,477.8 million PLN, with the amount of 20,831.8 million was spent on new fixed assets and modernization of existing ones. In the territorial arrangement functioned most micro the provinces: [9,12]

- Mazowieckie (16.4% of the entire population),
- Śląskie (11.6%),
- Wielkopolska (10.1%).

It may be noted that the distribution of value added is relatively constant for many years: micro enterprises - 10%; small - 7%; average - 10% of GDP.

Table 1. The share of SMEs in GDP in the period 2004-2008 (%)

	2004	**2005**	**2006**	**2007**	**2008**
SMEs in total	48,3	47,6	47,5	47,0	46,9
MICRO	30,9	31,4	30,9	30,2	29,8
SMALL	7,5	7,3	7,3	7,1	7,3
MEDIUM	9,9	8,8	9,3	9,7	9,8

Source :Raport o stanie sektora MSP w Polsce w latach , pod red. Wilmańska A. , PARP, Warszawa 2010, s. 28

Micro, small and medium-sized enterprises deserve to be called "the driving force" of the Polish economy. Although usually have short reach and low impact on the environment, in which they operate, their large number of suitable SME sector of great importance in shaping the economy. The contribution of SMEs in building gross value added is about 50% and the number of people working in this sector represents two thirds of all employed. In addition, the importance of micro, small and medium-sized economy proves their high potential to adapt to market conditions, flexibility and mobility, which are extremely important factors for stable, steady and rapid development of not only Polish, but European economy [6].

3. CLASSIFICATION STANDARDS – EXISTING STATE

Classification standards should have a continuity of work on their improvement, so they can always fit to dynamically changing environment. Standards should be updated and determined separately for each economy, industry and enterprise groups, as well as depend on the size of the companies. Currently, most of the standards and

classification, which the entrepreneurs use, has been created at least a decade ago. It becomes necessary to diagnose current needs of the economy and update the expressed values of indicators which describe the state of the financial condition of companies. Suitable modernization of standards will diagnose condition of SME sector. Standards should be accompanied by appropriate supporting materials to facilitate interpretation, which will provide comprehensive analysis. Although classification standards are created by large institutions and businesses for their own purposes (e.g. Main Statistical Office, the Polish Banks Association, etc.), but they are not suitable for the SME sector. There is a need to assess the risk of performing business activities. [15] In Polish literature (Sierpińska, Ostrowska, Michalski, Marcinek, Kuziak) there are basic methods of financial risk assessment, however, there are no standards relating specifically to the SME sector.

Lack of adequate instruments for assessing the financial condition and risk contributes to a number of bankruptcies in the SME sector. Any entrepreneur wants to bankrupt, but numbers show how many of them collapses. Answer to the question of the very meaning of existence of micro many years ago gave Schumpeter, when formulated his famous thesis of "creative destruction". According to him, capitalism could not exist without the constant emergence of new companies on the ruins of those that have fallen. With the rotation of the economy is still subject to "reincarnation" is in a constant process of transformation, leading it to even higher levels. Even so, making the proper tool for the study assess the financial condition of the SME sector may result in no need for the use by SMEs of the "creative destruction".

4. FINANCIAL ANALYSIS OF THE SME SECTOR IND ITS ASSESSMENT-GENESIS

Financial Analysis historically is the oldest department of economic analysis. Initially, financial analysis had a narrow scope, it occurred only in practice and served mainly a documentary and explanatory function [4]. It explains only the numbers which appears in the balance sheet and the profit and loss account. Nowadays, the essence of the financial analysis is not just to count the financial indicators. It is an attempt to clarify the factors affecting the improvement or deterioration in performance of the company, is engaged in the assessment of the effectiveness of the company's activities, and its business risk [15]. The analysis should include only matters relevant to the organization, to give answers to questions, have a systemic nature, determine the causes and effects of the studied phenomena, explain deviations from the plans, objectives and standards, disclose barriers, spot unnoticeable so far irregularities, justify a diagnosis of the financial outcome. It should be objective, include a brief conclusion and information to support the decision-making process, as well as proposals for targeted action. "Improving" financial analysis, creating an

artificial image of the company, often aimed at calming the investors or hide excessive risk is short-sighted and does not serve the purpose of improving the situation of the company [4].

Financial analysis tools proposed in the literature are not addressed specifically to the SME sector, which, according to estimates of the European Commission in 2013 will account for approximately 99.7% of all registered businesses, of which 95.3% are micro enterprises. Despite the fact that there are numerous books and articles about financial analysis, only a few companies decide to carry out such analysis on their own. The reason is lack of information about the business entity and the lack of sufficient expertise of entrepreneurs, as well as the time needed for the preparation of financial analysis. We believe that the most important is the fact that there is shortcoming of appropriate tools and methods to assess the financial condition, which could be used by companies of the SME sector. Enterprises that take their knowledge from the literature, must make an independent selection of information gained in this way. It is not easy to make the correct choice of techniques and methods of analysis when there are various of possible books. In addition, not every solution presented by an author can be used in the practice of companies in the SME sector. Entrepreneurs on the Polish market, are increasingly interested in financial analysis and appropriate indicator levels. The accuracy of the financial evaluation may depend not only economic efficiency, but increase market share, increase competitiveness of the company, and sometimes even affect its survival. Importance of it's in the SME sector, where most businesses fail within the first 5 years of operation of the market is big. One of the important reasons for this may be wrong financial management. A tool to facilitate assessment of the company's financial situation is a set of financial ratios. These indicators are calculated mostly by comparisons over time, so by comparing the size obtained in three consecutive years (fiscal year and two years prior). This comparison is to determine only the direction and rate of change observed value. While they do not give the possibility to assess the position occupied by the company in the sector (industry). Only a comparison of the individual indicators with indicators obtained by the competitors creates the conditions for claiming a level of efficiency, and thus reasoning in terms of opportunities or threats to the continuation of the action [3]. In view of these circumstances, the Commission of Financial Analysis of the Scientific Council of the Association of Accountants in Poland in 2002 made an attempt to develop a set of indicators characterizing the financial position, financial cover assets and profitability of companies and has proposed a methodology for the calculation as well as calculating indicators for various sectors of the Polish economy. However, norms for mapping evaluation standards presented by the Commission are global values that are the result of quantitative statistical approximation of the level of this indicator globally expressed as the arithmetic mean. As standards are designed based on Central Statistical Office data, it can be assumed that large and medium-sized enterprises plays a significant role in its creation, and certainly mainly

companies which prepare financial statements. In conclusion, it can be argued that this methodology has significant limitations, in particular:

1. For each of the sampled research indicators are calculated by given formulas, based on a statistical picture contained in the F-02 financial statements. Small businesses, and especially micro-enterprises, as a rule, are not required to fill out the SP and F-02 form. In view of this a question presents itself: Is the obtained value is meaningful for entrepreneurs in the SME sector and whether the existing indicators from the literature are appropriate and possible to compare?
2. The results are obtained by removing outliers, which are eliminated from the sample. Further analysis covers only those indicators, which were located at designated intervals. Which is inconsistent with the risk assessment methodology.
3. The next step is to determine the basic descriptive statistics for the purified sample. This approach does not create a fully reliable picture of reality.

Proposed research is extremely important for several reasons. Firstly, knowledge about standards of financial condition assessment were based so far on a global basis industry indicators, which in turn are described quite extensively, not taking into account neither different types of enterprises nor their size. Flexibility of companies is a powerhouse for any free economy. In Poland, in the period after the political transformation in 1989, many people noticed (following the example of Western Europe) that there is needed to "to take the law into one's own hands" brings tangible, positive results. This means that companies in this sector have the ability to quick respond for changing needs and preferences of potential customers and can adapt to them. By entering markets with relatively low growth potential and filling market niches, they contribute significantly to raising the efficiency of the whole economy and are an important element of regional development. Development of the SME sector is determined by two types of factors: external factors relating to the environment and internal factors. From the perspective of the smallest enterprises these factors one should look at them as an assessment of barriers for further development.

Main types of barriers hindering the development of the SME sector are [5]:

- market and social barriers,
- financial barriers,
- barriers resulting from economic policies,
- legal barriers,
- informational barriers,
- barriers related to the state of infrastructure.

In the business environment, there are obviously opportunities for the development, expansion, taking advantage of them and it is conditional on overcoming barriers

previously described. Barriers appear in, but are not restricted to, such areas and phenomena, as:

- the nature of entrepreneurship,
- preferred ethical and moral patterns,
- social perception of micro-business,
- a particular form of competition in the economy,
- used sources of funding,
- fragmentation, type and scale of operation,
- the failure of the legal system,
- low efficiency of policies to support SMEs,
- weakness of representing the interests of SMEs,
- inefficient management of the financial aspects of business [8,13,14].

Entrepreneurs knows about the existence of terms like "profitability" and "liquidity", but they cannot affect on their value in practice. Era of crisis shoved how common it is. Enterprises in order to save its financial condition take risky steps to secure profit, for example: buying currency options or take out usurious loans [7].

Due to the lack of experience and a high risk of such operations, usually they only accelerate the deterioration of the financial situation of the company. Therefore, the identification of factors that determine the financial condition or the use of research methodology for measuring the impact of these factors in practice is still a up-to-date topic. It is difficult to agree with one standard, such as liquidity ratios, which will be matched to different industries, different types of businesses of all sizes. Therefore, it is reasonable theoretically and practically useful to develop classification standards for assessment of the financial condition of the SME sector, from the point of view of different sections, in particular because of the legal personality, size and industry, etc. The expected effects of research are standards - template of classification standards for financial condition evaluation for SMEs in defined (using the theory of grouping) groups of companies. It will be important also to propose measures to assess the financial condition relevant to the SME sector. Measures, that with limited financial data for the group of companies, will be able to calculate and compare to developed standards. In scientific aspect this project is a way to enable the development of standards for the assessment of the financial condition of the various cross-sections.

In utilitarian aspect empirical research and work carried out within the project will allow to create classification standards for financial evaluation of different types of enterprises, in order to clarify domain knowledge and provide tools in the form of a structured theoretical assumptions that can be used in business practice for effective financial management of enterprises.

Financial analysis tools proposed in the literature are not addressed specifically to the SME sector. Despite the fact that there are numerous books and articles about financial analysis, only a few companies decide to carry out such analysis on their own, mainly because of lack of synthetic information about the business entity and the

lack of sufficient expertise of entrepreneurs, as well as the time needed for the preparation of financial analysis. The most important is the fact that there is shortcoming of appropriate tools and methods to assess the financial condition, which could be used by companies of the SME sector. Available standards of financial condition assessment are based on global industry indicators, which in turn are described quite extensively, not taking into account neither different types of enterprises nor their size. Developing classification standards for assessment of the financial condition of the SME sector in the various sections will require carrying out quantitative research (showing the extent of the phenomenon) and qualitative (being its characteristics). The research will be supplemented by reference query and analysis of the phenomenon of "desk research" having a descriptive and comparative characteristic, which will focus of evaluation standards financial condition in point of view of different sections, in particular because of the legal personality, size and industry, etc.

Therefore, this requires different approach to various groups of enterprises, hence the use of specific standards of analysis based on which the business could be evaluated. In essence, the project will have to implement the stated objectives:

1. Demonstrate that in the case of the SME sector it is not sufficient to use classical standards to assess the financial condition.
2. Development of a methodology for estimating the classification standards in use by grouping (cluster analysis) of the SME sector.
3. The use of the theory of diagnosis to develop and test methods for determining the financial risks based on defined standards

5. THE PURPOSE AND JUSTIFICATION OF RESEARCH - SUMMARY

Conducted considerations reaffirm the need for further research on this topic. There is and response to the question posed in the title: *Assessment of the financial condition of SMEs: is there a need to establish new compulsory or voluntary financial standards?*

To the developing classification standards for assessment of the financial condition of the SME sector in the various sections will require carrying out quantitative research (showing the extent of the phenomenon) and qualitative (being its characteristics). The research will be supplemented by reference query and analysis of the phenomenon of "desk research" having a descriptive and comparative characteristic, which will focus of evaluation standards financial condition. Defined research tasks require the implementation of the following research plan:

1. Analysis, diagnosis and assessment of existing classification standards assess the financial condition of companies in Poland and over the world in the context of taking into account the diversity of economic enterprises.

2. Evaluation of effectiveness of industry standards usage in the assessment of the financial condition of companies in the SME sector.
3. Identification and analysis of the factors determining the evaluation of the financial condition of companies.
4. Determining the impact of factors on the financial condition of companies.
5. The analysis of the differentiating force for the indicators used in the study of the financial condition of companies in the SME sector - analysis of test results, including factor analysis with using computer programs.
6. Classification of the SME sector entities, taking into account different indexing criteria.
7. Selection and development of new measures that could constitute the basis for the assessment of the financial condition of the SME sector, in particular for small and micro enterprises.
8. Construction of classification standards for assessment of the financial condition that takes into account the classification distribution of companies.
9. Designing a model for the assessment of the financial condition and financial risk SME sector including the division of business entities classification (through identification with the use of cluster theory).

The effect of the research will be the development of multidimensional classification standards for the assessment of the financial health of the SME sector. Selected companies decide to conduct the assessment of the financial condition of the company on their own, because of lack of sufficient specialized knowledge among entrepreneurs, the time required for the absorption of it and the financial information necessary to complete it. The most important is the lack of adequate methods for assessing the condition of the SME sector based on measures relevant to the group of companies which respects the limitations in the source data. The standards that will be developed as a result of this project will be a valuable and relatively cheap source of information, and also will allow businesses to manage economic efficiency and make rational decisions.

REFERENCES

[1] Baranowska-Skimina A., *Sektor MSP w Polsce w latach 2008-2009*, http://www.egospodarka.pl/, dn. 21.03.2011 r.
[2] Bukowski M., *Małe i średnie przedsiębiorstwa złagodziły skutki kryzysu*, Deutsche Bank 13.08.2010 r.,
[3] Dudycz, M. Hamrol, W. Skoczylas, A. Germany, 2005, *Financial sector indicators - help at financial analysis and evaluation of the company's ability to continue operations,* Accounting 3/2005
[4] Gabrusewicz W., *Fundamentals of financial analysis*, PWE, Warszawa 2005, p. 17
[5] Grzybowska B., Juchniewicz M., *„Innowacyjność mikroprzedsiębiorstw w Polsce”,* Polska Agencja Rozwoju Przedsiębiorczości, Warszawa 2010, pp. 8–11.

[6] Informacja o handlu zagranicznym Polski na koniec grudnia 2010 r., Departament Analiz i Prognoz przy Ministerstwie Gospodarki, Warszawa, luty 2011 r., p. 3.
[7] Jacs Z., Audretsch D. B., *Small firms and entre-preneurship: an East-West perspective*, Cambridge University Press, 1993, s.1, via B. Mikołajczyk, M. Krawczyk, *Sektor przedsiębiorstw mikro, małych i średnich w krajach Unii Europejskiej*, Szkoła Główna Handlowa, Warszawa 2006.
[8] *„Mikroprzedsiębiorczość w Polsce"* research conducted by Pentor Research International or Kronenberg Foundation in Citi Handlowy with cooperation with Microfinance Centre
[9] Wilmańska A. *Raport o stanie sektora MSP w Polsce w latach*, , PARP, Warszawa 2010, p. 28
[10] Siemińska E., *Metody pomiaru i oceny kondycji finansowej przedsiębiorstwa*, Uniwersytet Mikołaja Kopernika w Toruniu, Toruń 2002
[11] Skowronek-Mielczarek A., *Małe i średnie przedsiębiorstwa. Źródła finansowania,* Warszawa 2003, p. 7.
[12] Wilmańska A., *Stan i tendencje rozwojowe sektora MSP w Polsce w latach 2008-2009, Materiały konferencyjne PARP,* Warszawa 20 paź. 2010 r., p. 22.
[13] Źródło: http://pkpplewiatan.pl/dla_mediow/informacje_prasowe/1/_files/2011_12/Raport_MSP.pdf
[14] Skoczylas W., *Analiza wskaźnikowa w zarządzaniu przedsiębiorstwem, Prace i materiały UG, Współczesne problemy analizy ekonomicznej,* Sopot 2006

Web browsers, HTML5,
Web content delivery latency,
Web page processing, QoE

Mariusz FRAŚ,
Ziemowit NOWAK*

ON SOME ISSUES OF WEB CONTENT PROCESSING EFFICIENCY ON THE END USER SIDE

The research on the quality of Web services usually take into account the network latency and the server-side latency. The chapter concerns selected topics of quality of Web content delivery on the end user side. The contemporary Web services involve not only processing of user requests on the Web server, but also produce processing by user agents (usually Web browsers). The work presents general schema of content processing by user agents, and some issues of Web transaction latency measurement. Then, some results of research on efficiency of three most popular Web browsers, and the impact of selected parameters of delivered content on client-side latency is shown. The presented results of experiment mainly concern the processing time of new HTML5 elements.

1. INTRODUCTION

The quality of Web content delivery has been investigated for several years, and is very current area of interest in computer science. Especially performance and content delivery latency from the end user point of view is still explored problem motivated by constant development of Web technologies.

The general scheme of Web transactions usually takes into consideration [e.g. 1, 2, 3] the phase of establishment of connection with the server (time of choosing the service instance T_{DM}, resolving DNS name time T_{DNS}, and TCP Connect time T_{TCPC}), request transfer, request processing on the server, and transfer of the results to the client (fig. 1.). This means that network latency and request processing on the server latency is taken into account.

* Wroclaw University of Technology, 50-370 Wroclaw, Wybrzeze Wyspianskiego 27.

Many works considered different approaches to improve Web service quality such as choosing the "nearest" Web server [e.g. 4, 5], Internet distance metric estimation [e.g. 6, 7], local and global client request distribution [e.g. 8, 9], and many others.

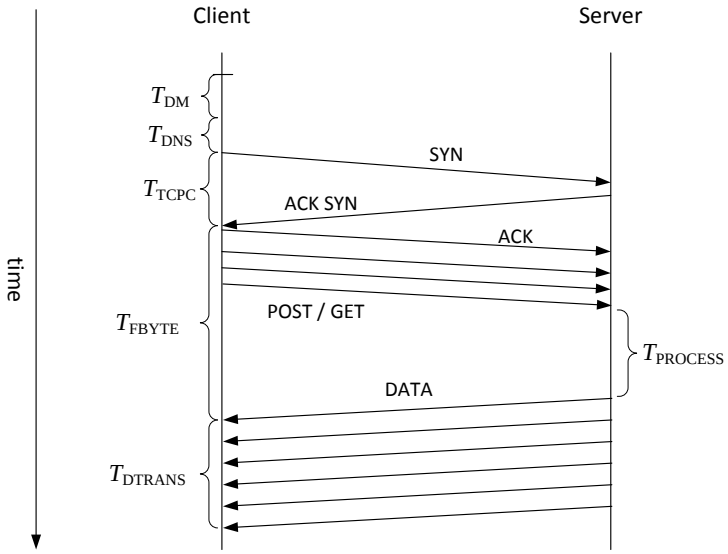

Fig. 1. The Web transaction – the TCP level view

However, the analysis of network and server latency only, refers well to the case of static Web content or creating dynamic content only on the server. From several years the Web becomes more and more dynamic, multimedia and interactive. The end user activity involves not only processing of client request on the Web server, but also produces processing (sometimes quite advanced) by user agents (usually Web browsers). While the network communication efficiency constantly increases, the latency of the end user agent processing becomes important part of Web service Quality of User Experience (QoE).

This chapter considers selected topics of Web content processing efficiency on the end user side. In addition it takes into account new mechanisms of HTML5 standard and compares some of them to other one (namely Flash). First, the general schema of content processing by user agents and latency measurement issues are presented. Next, some results of research on selected issues of user agent latency are shown. The measurement experiment concerned the processing time of new HTML5 Web content elements on client side (especially video and animation). Experiment compares efficiency of three most popular Web browsers, and shows the impact of selected parameters of delivered content on client-side latency.

2. USER AGENT-SIDE WEB CONTENT PROCESSING AND MEASUREMENT

2.1. PROCESSING HTML DOCUMENTS

The HTML documents have the structure of a tree of elements defined in the HTML specification. The specification defines also rules how elements can be nested, and the attributes, which control how they work. HTML5 user agents, usually Web browsers, parse HTML5 documents, and turn it into a DOM (Document Object Model) tree. A DOM tree is representation of parsed document in device's memory (e.g. computer RAM). DOM trees contain a number of nodes. A DOM tree can be manipulated from scripts (usually JavaScript) in the Web page. DOM trees are processed and presented by user agents on the end user side – HTML documents are rendered to a screen.

The loading of Web page in contemporary Web browsers is very complex process. The W3C (World Wide Web Consortium) specification [10] describes the Web application architecture and gives default rules and mechanisms how Web browsers should work. In a HTML browser, every HTML document is represented by a *Document* object. *Document* objects are presented to the user in so called browsing context. A browsing context has a session history – a list of objects that has been presented, and an active document. In general, browsing context processing is very complex. It involves among others nested contexts, groupings of contexts, security, sandboxing, etc. There are several models of page load processing, e.g. for HTML files, for XML files, for content that uses plugins etc. For HTML files when an HTML document is to be loaded in a browsing context a *Document* object is created, its content type is set to "text/html", and an HTML parser is created and associated with the document. All bytes fetched from content source (or sources) are passed to the parser which process them. When the stream of bytes ends a load event is fired.

The figure 2. shows the parsing process of HTML content. The input consists of (a little simplifying) a stream of Unicode code. It is passed to a Tokenizer. The output of the tokenization step is a series of tokens of special format used to construct the DOM tree. When a token is emitted, it handled by the tree construction stage which dynamically modifies or extends document's DOM tree (it can create and insert nodes in DOM tree). The output of parsing process is a *Document* object.

When the browser stops parsing the document, it runs a set of steps that finishes with queue a task to mark the *Document* as completely loaded.

As we can see, the processing performed by user agents can be complex and take noticeably long time, for example due to script processing. As stated in [10] "parsing of HTML files happens asynchronously and incrementally, meaning that the parser can pause at any point to let scripts run". Moreover: "scripts in HTML have "run-to-

completion" semantics, meaning that the browser will generally run the script uninterrupted before doing anything else, such as firing further events or continuing to parse the document". All this can significantly impact the user latency which is very important quality attribute for Web content delivery.

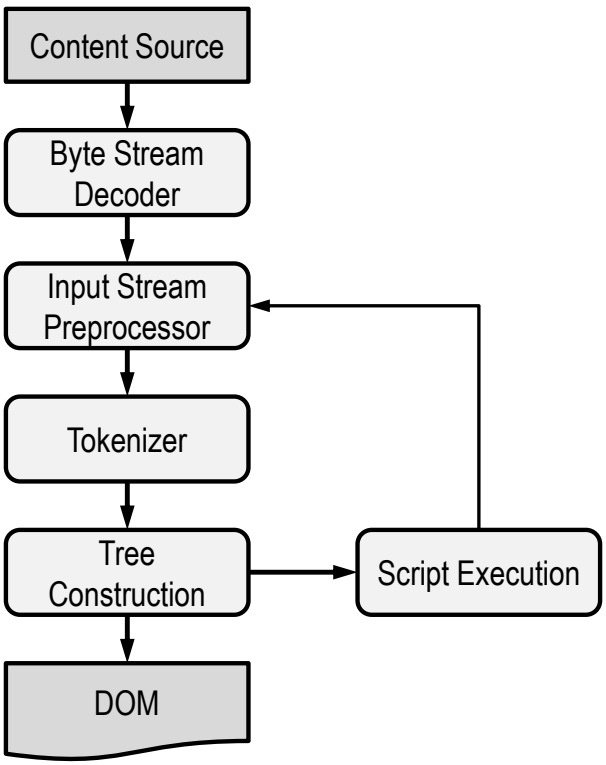

Fig. 2. The general schema of the HTML parsing process (on the basis of [11])

2.2. THE NAVIGATION TIMING API

To get detailed timing data for page navigation, rendering and loading the W3C developed specification Navigation Timing [11] which defines interfaces for Web applications to record detailed timing-related information. Navigation Timing is a JavaScript API which permits to get detailed timing statistics without affecting process of Web page loading. The specification defines two interfaces (PerformanceTiming interface and the PerformanceNavigation interface), and extends the third one (HTML5 Window interface) that specify timing attributes which recording should be supported by user agents (Navigation Timing API should be directly built into the user agents). The figure 3. shows a little simplifying case (the redirection and unloading previous page is omitted) of registered timing statistics.

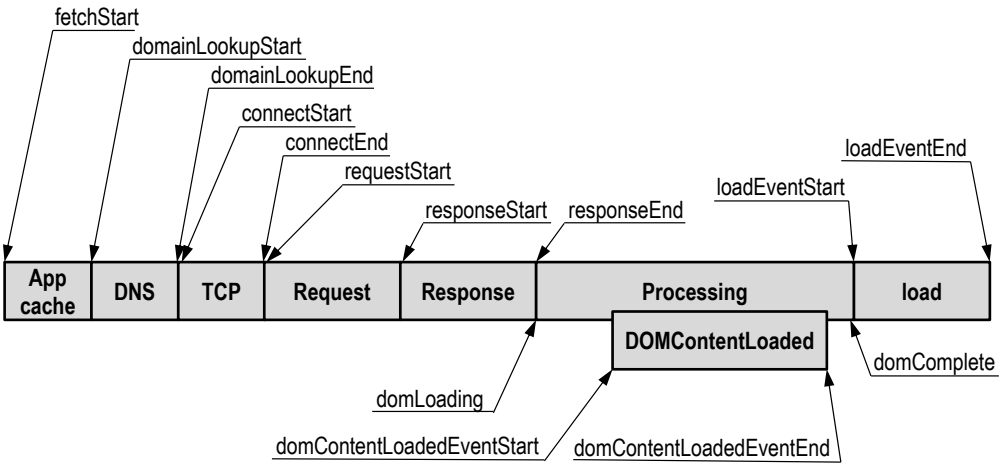

Fig. 3. The timing attributes for simplified Web navigation defined in Navigation Timing specification

The recorded times can be used to measure DNS lookup time, TCP Connect time, the time of client request and server response transfer time as well as the processing and loading time in user agent, including the time of building DOM tree and script processing. Thus, e.g. the network latency with included server latency (i.e. responseEnd-fetchStart), and the time taken for page load after it is received from the server (i.e. loadEventEnd-responseEnd) can be measured.

At the present, three most popular Web browser, i.e. Firefox (starting from ver. 7), Internet Explorer (starting from ver. 9), and Chrome (starting from ver. 6) support Navigation Timing API. This gives a standard way of measurement of performance-related data for Web content delivery also on the end user side.

3. RESULTS OF EXPERIMENT ON BROWSER PROCESSING EFFECTIVENESS

3.1. THE OBJECTIVES OF PERFORMED EXPERIMENTS

The goal of performed experiments was to test new features of HTML5 in the context of browsers processing effectiveness and comparison HTML5 to other technologies. By browser processing effectiveness we consider here the Navigation Timing processing time plus load time of Web page (see figure 3.).

In every measurement three Browsers which implement Navigation Timing API, were tested: Firefox 12.0, Chrome 19.0.1084.52, and IE 9.0.8112.16421. The two HTML5 new elements were chosen for tests: the Video element and Canvas element.

The Video element is very useful because eliminates the need to install plug-ins for video playback (e.g. Flash Player), and provide a uniform way of presenting movies by Web browsers. Such an approach makes it possible to apply CSS styles and JavaScript, and permits individually customize the look of the Player and its available features on the website. The standard don't specify exact video formats, and different browsers support different sets of formats. Therefore, when delivering video in HTML5 several versions should be provided on the server. The experiment was aimed to compare effectiveness of processing pages with Video element, and the impact of various encoding on processing effectiveness.

The Canvas element is the large part of the HTML5 specification. It is used for animations, for interactive backgrounds of Web pages, for navigation, as graphical tools, as independent web applications, and other. Currently, all major browsers support the Canvas element. His strength in relation to Flash, which is used for similar applications, is that the graphics works even on mobile devices, which cannot be said about the pages in Flash [12]. The second objective was to compare effectiveness of browsers processing of Flash animations, and HTML5 animations coded from scratch and converted from Flash.

3.2. BASICS OF MEASUREMENT PROCEDURE

Since many web sites delivers video content of approximately 30s length, four differently encoded movie copies were set up. The following video formats were used: WEBM, OGV, MP4, and FLV.

For Canvas and Flash comparison there were 14 different animation prepared – 7 for comparison of Flash, and converted from FLV, HTML5 animations, and 7 for comparison of Flash and native (coded from scratch) HTML5 animations. For the first group, seven FLA source files (Flash sources) were used, and with Adobe Flash Professional CS6 seven SWF files and seven HTML5 animations were prepared. It must be noticed here that generated this way the animation code for the HTML5 Canvas is not optimal. For the second group animations were created independently.

Video and Canvas Elements were embedded on separate web pages and placed on a web server. Each web page was downloaded 10 times and the timing statistics were recorded on the client side. Especially the difference loadEventEnd–domLoading was measured. Then the median was calculated. To prevent unwanted impact on measurements, before page downloading the cache was cleaned and removed Cookies.

3.3. PROCESSING TIME DEPENDING ON THE MOVIE FORMAT

In figure 4. the processing time depending on the used video format is shown. For Internet Explorer the results for all formats are at the same level. The similar situation

occurs in the case of Firefox. It appears that for these browsers presented movie format does not matter, and in any case it is better than the Flash (FLV). For Chrome the times distinctly differ and are clearly longer. It is caused by the necessity of waiting for loading initial seconds of the movie. Other browsers load video only when you press "play". Please note that OGV format is not supported by Chrome (and the measured time is the lowest for Chrome). Chrome works best with WebM format. Despite the fact that Firefox does not support MP4, and Internet Explorer - OGV, timings for all formats are similar.

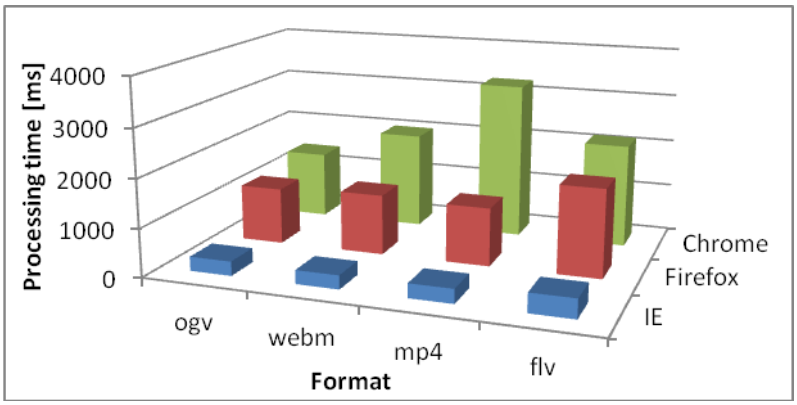

Fig. 4. The processing time of page with Video element depending on movie format

The results show that the two browsers for the movie format does not affect the time of page processing. It seems that IE is clearly the most effective browser; however it must be emphasized the fact of preloading the movie by Chrome and testing old version of Firefox. The newer versions are more efficient. Flash format (FLV) is for every browser slower. This is probably due to the fact of loading the appropriate plugin.

3.4. PROCESSING TIME FOR HTML5 CANVAS AND FLASH ANIMATIONS

In the beginning it must be recalled that 14 different animations were used for measurements and the presented results should be interpreted mainly for comparison of Flash and Canvas each other.

The figure 5. shows page processing time depending on the used technology – Flash and Canvas, wherein HTML5 code was obtained from FLA source. It can be observed for Firefox that for four HTML5 animations processing time is longer, and for three ones is shorter, and the maximum difference between these times is 300 ms. For Chrome the inverse proportion occurs, but also can be seen that the times are comparable. Exactly the same phenomenon occurred for Internet Explorer. It must be

noticed here that conversion from FLA to HTML5 code makes Canvas animation probably less efficient that making the code from scratch.

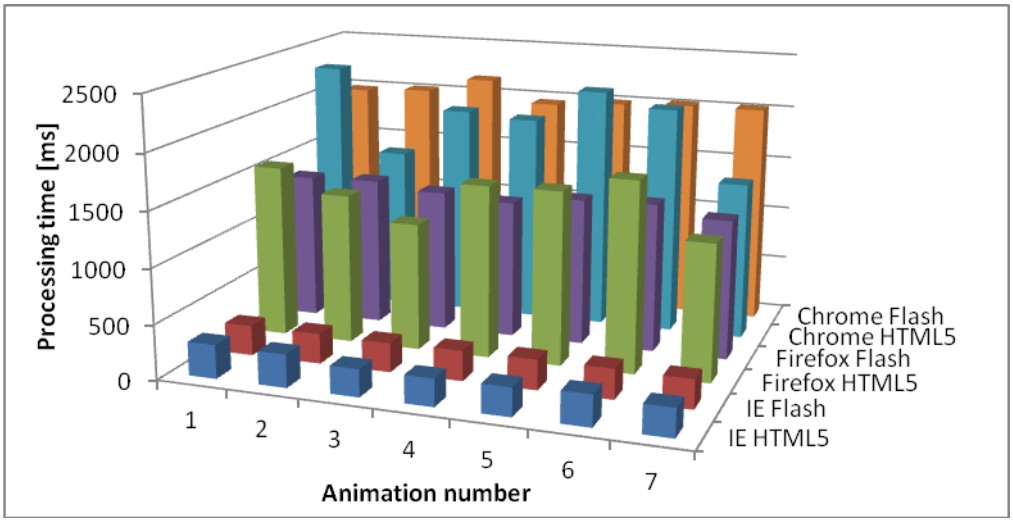

Fig. 5. The processing time of page with Flash and HTML5 Canvas (converted from FLA)

The figure 6. shows page processing time depending on the used technology – Flash and Canvas, wherein HTML5 was coded from scratch. It must be noted again that measurements shows only corresponding relationships, when the code was written from scratch. The content of the animation was not the same for all seven cases.

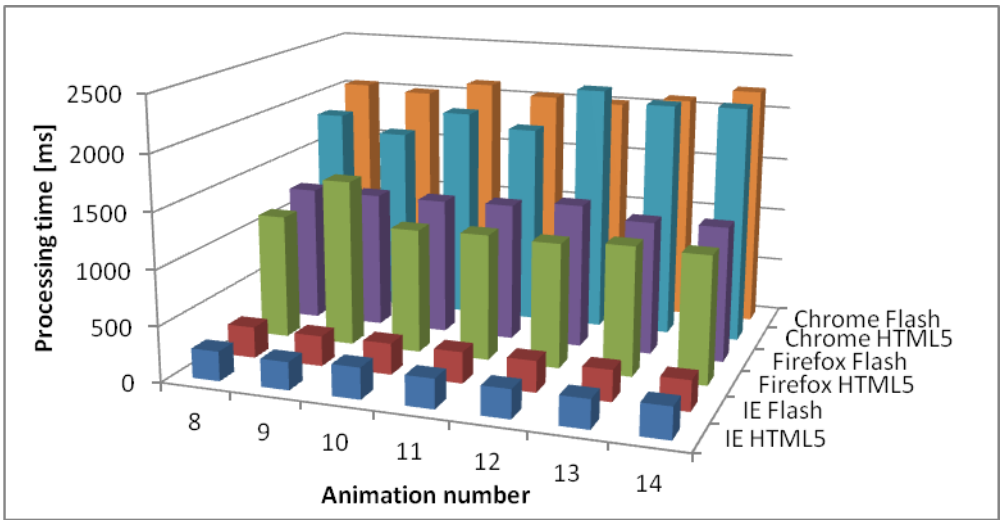

Fig. 6. The processing time of page with Flash and HTML5 Canvas (native)

For this experiment it can be seen that for Firefox and Chrome HTML5 gives generally better results than Flash and the differences of effectiveness between these technologies are not significant.

FINAL REMARKS

Navigation Timing provides useful tools for detailed measurement of Web browsing timing statistics. In presented studies only a fraction of the capabilities of the Navigation Timing API was used. To make necessary measurements only two variables defined in the interface were used: loadEventEnd and domLoading. For example, very useful can be properties concerning heap usage, i.e.: totalJsHeapSize, jsHeapSizeLimit or usedJsHeapSize, and many more.

Presented results generally compare three most popular web browsers. However researcher must remember that the devil is in the details. In this work the fact of preloading first seconds of movie in Chrome is an example. Also surprisingly significant better efficiency of IE should be a lead to explore deeper this behavior. Some characteristics suggest that IE may not really measure the same as other browsers.

On the basis of Navigation Timing API you can build a desktop program or a web application, which allow configuring parameters and downloading different web pages for research. The application could also contain a module for processing the generated data and presenting them on the charts. Such software could also present other measured parameters, offered by API. Better reporting helps to understand page load latency. That allows building faster web applications, and better websites.

ACKNOWLEDGMENTS

The authors thank to mgr inż. Jakub Cieślar for the preparation and carrying out experiments.

REFERENCES

[1] BORZEMSKI L., NOWAK Z., *WING: a web probing, visualization, and performance analysis service*, in Proc. Web engineering. 4th International Conference (ICWE 2004), Munich, Springer, 2004.

[2] BORZEMSKI L., CICHOCKI Ł., FRAŚ M., KLIBER M., NOWAK Z., *MWING: a multiagent system for Web site measurements*, Lecture Notes in Computer Science, Lecture Notes in Artificial Intelligence, vol. 4496, 2007.

[3] FRAŚ M., KWIATKOWSKI J., *Quality aware virtual service delivery system*, Journal of Telecommunications and Information Technology, vol. 3, Instytut Łączności - Państwowy Instytut Badawczy, Warszawa, 2013.

[4] BORZEMSKI L., NOWAK Z.: *Using the Geographic Distance for Selecting the Nearest Agent in Intermediary-Based Access to Internet Resources*. LNAI Vol. 3683, Springer-Verlag, Berlin, 2005.
[5] BORZEMSKI L., NOWAK Z., *Analiza skuteczności metod wyboru najbliższego sąsiada w Internecie*, Współczesne aspekty sieci komputerowych, Wydawnictwa Komunikacji i Łączności, Warszawa, 2008
[6] FRAŚ M. GIEDROJĆ G., *IP path clustering-based distance metric*, Information systems architecture and technology : new developments in Web-Age Information Systems / eds. Leszek Borzemski [i in.], Oficyna Wydawnicza Politechniki Wrocławskiej, 2010.
[7] ANDRZEJCZAK S., NOWAK Z., *Metody wyznaczania dystansu w Internecie*, Współczesna problematyka sieci komputerowych, Wydawnictwa Komunikacji i Łączności, Warszawa, 2010.
[8] BORZEMSKI L, ZATWARNICKA A., ZATWARNICKI K., *Global distribution of HTTP requests using the fuzzy-neural decision-making mechanism*, Proc. of 1st Int. Conf. on Comp. Collective Intelligence, In Lecture Notes in AI, Springer, 2009.
[9] FRAŚ M., ZATWARNICKA A., ZATWARNICKI K., *Fuzzy-neural controller in service request distribution broker for SOA-based systems*. In Proc. of Int. Conf. Computer Networks 2010, Kwiecien A., Gaj P., Stera P. (eds), Springer, Berlin, Heidelberg, 2010.
[10] BERJON R., FAULKNER S., LEITHEAD T., et al, *HTML5: A vocabulary and associated APIs for HTML and XHTML*, W3C Recommendations, World Wide Web Consortium, 2014, http://www.w3.org/TR/html5/
[11] WANG Z., *Navigation Timing, W3C Recommendations*, World Wide Web Consortium, 2012, http://www.w3.org/TR/navigation-timing/
[12] LAWSON B., SHARP R., *Introducting HTML5 Second Edition*, New Riders, Berkeley, 2012

PART 2

RISK MANAGEMENT IN DECISION MAKING PROCESS

Project portfolio selection, fuzzy simulation
probability, possibility, dependence

Bogdan RĘBIASZ,
Bartłomiej GAWEŁ,
Iwona SKALNA*

CAPITAL BUDGETING OF INTERDEPENDENT PROJECTS WITH FUZZINESS AND RANDOMNESS

This paper discusses the problem of capital budgeting in the situation where only some model parameters are well described by their past data and therefore are specified by random variables, whereas the remaining model parameters can hardly be predicted by historical data and therefore they are described by means of fuzzy variables. In order to be able to process such hybrid data, a model of the problem is proposed. The model takes into account both stochastic and economic interdependency between projects. Additionally, a new hybrid method for solving this model is developed. The method combines stochastic simulation with arithmetic on interactive fuzzy numbers and nonlinear programming. As a result a set of Pareto-optimal alternatives is obtained. In order to illustrate the performance of the proposed hybrid method, an example from metallurgical industry is provided.

1. INTRODUCTION

Accurate prediction of investment outlays is very difficult in an uncertain economic environment. Therefore, decisions related to the realization of investment projects are unmanageable without appropriate description of investments and identification of financial and material problems. The allocation of a company's capital to a combination of investment projects, which brings the company a maximal total profit is referred to as the *capital budgeting problem* (*CBP*) [5].

An important aspect of the capital budgeting problem is the analysis of dependency, which is characteristic for many economic problems. In the frames of the *CBP*, generally the two kinds of dependency can be distinguished, the one between model parameters and the one between projects. The economic interdependence

* AGH University of Science and Technology, Kraków, Poland.

between projects shows that one project is influencing the effects generated by other projects. Whereas, the statistical dependency is used to measure the strength of dependencies (associations) between probability distributions or fuzzy variables [9,11].

This paper present a solution to the capital budgeting problem, where projects are statistically and economically dependent. It is assumed that uncertainty attached to model parameters origins from two sources: randomness and imprecision. Distinct approaches are used to adequately represent these two kinds of uncertainty. The objective uncertainty (randomness, variability) is represented by a probability distribution, whereas the subjective uncertainty (imprecision) is described by a possibility distribution (fuzzy numbers). In order to be able to jointly propagate both kinds of uncertainty a new procedure is proposed. The procedure also enables to take into account the two types of dependencies listed above.

The remaining part of the paper is organized as follows. Section 2 presents a capital budgeting problem (*CBP*). In Section 3, a hybrid method for solving the *CBP* is described. Section 4 presents a numerical example that illustrates the performance of the proposed hybrid method.

2. THE CAPITAL BUDGETING PROBLEM

2.1. INTRODUCTION

The problem of building an effective capital budget is divided here into two models – the *portfolio selection model* (*PSM*) and *portfolio evaluation model* (*PEM*). The purpose of first model is to find an efficient portfolio of investment projects. Whereas the second model is used to determine efficiency parameters for the set of investment projects obtained from the *PSM* model. The *PSM* is a multi-criteria binary linear programming model, while *PEM* is a non-linear programming model combined with stochastic simulation.

2.2. PORTFOLIO SELECTION MODEL

Let us consider a company in which there are m potential interdependent projects available for realization. Let us also assume, that each project creates new or modifies existing process steps within a primary production process. A portfolio of investment projects is defined as $(y_1,\ldots,y_w)$, where $y_i = 1$ if project i ($i \in W=\{1,\ldots,m\}$) is selected for realization, and 0 otherwise. Let $fin(y_1,\ldots,y_w)$ denotes a financial evaluation parameter for a given portfolio of investments. The performance of the selected portfolio is measured by the expected value $E(fin(y_1,\ldots,y_w))$ and the standard deviation

$\sigma(fin(y_1,\ldots,y_w))$. Then, the problem of selecting the efficient portfolio of investment projects can be defined as follows.

Find $(y_1,\ldots,y_w)$ that maximizes the expected value of $fin(y_1,\ldots,y_w)$ and minimizes the standard deviation of $fin(y_1,\ldots,y_w)$.

In order to solve the *PSM* problem, the *PEM* model must be first invoked for each portfolio of investment projects in order to obtain the respective value of the financial evaluation parameter.

2.3. PORTFOLIO EVALUATION MODEL

The value of financial evaluation parameter of a portfolio of investment projects is estimated by solving a model which consists of two groups of equations. The first group describes balances of a company's manufacturing capacities and material balances, and the second one consists of financial equations.

Equations of manufacturing capacities balance for primary production process are the following:

$$\sum_{i\in I} X_{ijw}^{t\tau} \le v_{jw}^{\varsigma}\cdot \Delta_{jw}^{\tau} \text{ for } \tau=0,1\cdots\bar{\tau},\tau\le t, j\in J, w\in W_j, t=\tau,\tau+1,\cdots,\tau+\bar{t}_{jw} \tag{1}$$

$$X_{ijw}^{t\tau}\ge 0\,,\ \varsigma=t-\tau\,, \tag{2}$$

$$\Delta_{jw}^{\tau}=\begin{cases}1, & \text{for } w\in\bar{W}\\ 0, & \text{for } w\in W-\bar{W}\end{cases}, \tag{3}$$

$$\kappa\left(\bar{W}\right)=1, \tag{4}$$

$$\eta^{\tau}\left(\bar{W}\right)\le\bar{\eta}^{\tau}\text{, for } \tau=0,1,\cdots,\bar{\tau}\ , \tag{5}$$

whereas the equations for the enterprise material balance are as follows:

$$\sum_{j\in J}\sum_{w\in W_j}\sum_{\tau=1;\tau\le t}^{\bar{\tau}} X_{ijw}^{t,\tau}-\sum_{j\in J}\sum_{w\in W_j}\sum_{z\in \mathrm{I}}\sum_{\tau=1;\tau\le t}^{\bar{\tau}} m_{izjw}X_{zjw}^{t,\tau}=G_i^t \quad \text{for} \quad t=0,1,2,\cdots,T, \tag{6}$$

$$G_i^t\le\bar{g}_i^t\left(\bar{W}\right) \quad \text{for} \quad t=0,1,2,\cdots,T, \tag{7}$$

where:

$X_{ijw}^{t\tau}$ – variable determining the quantity of the gross output of the product *i* produced in the department *j* in year *t*, in case of qualifying the project *w* for the realization in year τ,

$\bar{\tau}$ – capital budgeting period

T – time horizon of the optimization

G_i^t – variable determining the size of sale of the product *i* in year *t*,

I – set of product indexes,

I_j – set of indexes of products produced in the department *j*,

W – set of project indexes,

W_j – set of indexes of projects connected with the department *j*,

$\overline{W}$ – set of indexes of projects qualified to realization,

J – set of primary production department indexes

v_{jw}^{ς} – manufacturing capacity of the department *j* after realization of the project *w* in ς year of the duration,

$\bar{\eta}^{\tau}$ – limit of investment outlays in the year τ,

m_{izjw} – consumption per unit of the product *i* for producing the product *z* in the department *j* after realizing the project *w*,

$\bar{t}_{jw}$ – duration of the project *w* being realized in the department *j*

c_i^t – selling price for the product *i* in year *t*

r_d – long-term interest rate

r_k – short-term interest rate

$\bar{\kappa}: 2^W \to \{0,1\}$ – function determining sets of projects being possible for the realization, value 1 means a set possible for the realization, value 0 means set impossible for the realization,

$\eta^{\tau}: 2^W \to R$ – function assigning to $\bar{W}$ set of the projects an investment outlay for realization of this set in τ year of capital budgeting period

$\bar{g}_i^t : 2^W \to R$ – function assigning to $\bar{W}$ set of the projects possible sale of the product in the *t* year

The set of financial equations express commonly known dependencies such as balance sheet, *P&L* account and net cash flows (*NCF*). Their detailed presentation would considerably increase the volume of the article, therefore, they are omitted.

These two groups of equations are constraints and the goal function is defined as the financial efficiency parameter *EffPar*. Usually, the net present value (*NPV*) is used as a financial evaluation parameter. In this article, a company's gross profit (*GP*) is used instead. This approach simplifies the model and allows us to focus on presenting a hybrid method for processing possibilistic and probabilistic variables. The company gross profit (*GP*) is presented below:

$$GP^t = \sum_{j\in J}\sum_{i\in \mathrm{I}_j} c_i^t G_i^t - \sum_{\tau=1,\tau\le t}^{\bar{\tau}} \sum_{w\in W_j}\sum_{j\in J}\sum_{i\in \mathrm{I}} k_{ijw}^{\varsigma} X_{ijw}^{t,\tau} - r_k STC^t - r_d LTC^t - \chi^t\left(\bar{W}\right) + \xi^t\left(\bar{W}\right) \qquad (8)$$

where:

STC^t – variable determining the value of short-term credit in year *t*,

LTC^t – variable determining the value of long-term credit in year *t*,

GP^t – variable determining the gross profit in year *t*,

k_{ijw}^{ς} – cost of processing the product *i* by the department *j* after realization of the project *w* in ς year of the duration

$\chi^t\left(\bar{W}\right)$ – function assigning a company's fixed costs without amortization in year *t* to a project portfolio

$\xi^t\left(\bar{W}\right)$ – function assigning the value of amortization in year *t* to a project portfolio

3. SOLUTION METHOD

3.1. DESCRIPTION OF UNCERTAINTY IN CAPITAL BUDGETING PROBLEM

Risk analysis recognizes two types of uncertainty – aleatory and epistemic. The aleatory uncertainty is due to variability or randomness, whereas the epistemic uncertainty comes from the ignorance or the lack of knowledge. In the case of economic calculus, data may come from a variety of sources, and therefore it is usually heterogeneous, i.e., both random and imprecise. The most common situation in practice is when for some parameters it is possible to determine probability distributions (there is a sufficient enough amount of historical data), while some information is available in the form of possibility distributions (obtained from subjective assessments of phenomena made by experts) [2,7,8]. These two methods for description of uncertainty of parameters of economic calculus are used usually as alternatives. There are few studies which describe the use of hybrid data – data partially described by probability distributions, and partially by possibility

distributions. The use of such (hybrid) data allows to reflect more properly the knowledge on parameters of economic calculus [1,3,4,7].

The most common framework for representing and reasoning with uncertain knowledge is the Dempster–Shafer (*D–S*) theory of evidence. The *D–S* theory allows to treat variability and imprecision together in single framework. This theory is based on *D–S* probability mass structures. A *D–S* structure is a mass function which is much the same like a discrete probability distribution except that probability is attached to a set of values (intervals) instead of single points. Dempster–Shafer theory is widely used by many authors (Guyonnet *et al*. [4], Ferson [3], Baudrit [1] Cooper *et al*. [11] and others) to develop techniques of uncertainty propagation. On the other hand, many authors present methods for calculating measures of risks (for example the standard deviation of *NPV*) on the basis of the *D–S* theory, random fuzzy set theory and credibility theory [6].

3.2. SOLUTION OF PORTFOLIO EVALUATION MODEL

To solve the Portfolio Evaluation Model (*PEM*) it has been decided to use the hybrid propagation method proposed by Baudritt *et al*. [1]. The hybrid propagation method combines a Monte Carlo simulation with the extension principle of the fuzzy set theory.

Consider the *PEM* model which measures efficiency of a given portfolio of projects $(y_1,\ldots,y_m)$. *PEM* is a linear programming model with p parameters. These parameters can be divided into two groups based on the type of uncertainty attached to them. The first group consists of probabilistic parameters and the second one of possibilistic parameters. In the model considered in this paper, probability distributions are used to describe demand and selling prices, whereas material consumption and product prices are described by possibilistic variables.

The overall computational procedure can be summarized as follows. In order to solve the *PEM* model, possibility distributions are first divided into a set of intervals using the α-cuts approach. For each interval (α-level) a Monte Carlo sampling of random variables is performed taking into account dependencies between those variables. The dependencies are processed using a method presented by Yang [10], which employs the Cholesky decomposition of the correlation matrix. The Cholesky decomposition is commonly used in Monte Carlo-based methods for simulating systems with multiple correlated variables. Obtained realizations of random variables are put into the *PEM* model. Moreover, for each α-level, interval parameters are converted into variables in the *PEM* model and the following constraint is added to the model for each converted parameter:

$$\inf\left(\tilde{\boldsymbol{X}}_i\right)_\alpha \le x_i \le \sup(\tilde{\boldsymbol{X}}_i)_\alpha, \tag{9}$$

where $\inf\left(\tilde{\boldsymbol{X}}_i\right)_\alpha$, $\sup(\tilde{\boldsymbol{X}}_i)_\alpha$ are respectively lower and upper bounds of the respective α-level of a fuzzy parameter $\tilde{\boldsymbol{X}}_i$.

Dependencies between fuzzy parameters are described by means of interval regression. Interval regression parameters are reflected by two additional constraints:

$$\boldsymbol{x}_i \geq \inf\left(a_1^{iz}\right)\cdot \boldsymbol{x}_z + \inf\left(a_2^{iz}\right), \tag{10}$$

$$\boldsymbol{x}_i \leq \sup\left(a_1^{iz}\right)\cdot \boldsymbol{x}_z + \sup\left(a_2^{iz}\right), \tag{11}$$

where $\sup\left(a_1^{iz}\right), \inf\left(a_1^{iz}\right), \sup\left(a_2^{iz}\right), \inf\left(a_2^{iz}\right)$ are respectively lower and upper bounds of the interval regression coefficients describing the dependency between parameters $\tilde{\boldsymbol{X}}_z$ and $\tilde{\boldsymbol{X}}_i$.

Next, in order to determine the lower and upper bounds of the respective α-level of the financial efficiency parameter, the following constrained optimization problems must be solved:

$$EffPar_\alpha \rightarrow \min \tag{12}$$

for the definition of the lower bound of the α-level of the *EffPar*, and

$$EffPar_\alpha \rightarrow \max \tag{13}$$

for the definition of the upper bound of the α-level of the *EffPar*.

The optimization problems (12) and (13) are solved for each α-cut and as a result a fuzzy number μ^{EffPar} is found. Drawing probabilistic values and determining a realization of μ^{EffPar} is repeated n times. As a result, n a family of fuzzy numbers ($\mu_1^{EffPar},\cdots,\mu_n^{EffPar}$) are obtained. Based on the vector ($\mu_1^{EffPar},\cdots,\mu_n^{EffPar}$), the mean value and standard deviation of $fin(y_1,\ldots,y_w)$ are calculated.

3.3. SOLUTION OF PORTFOLIO SELECTION MODEL

The hybrid method for uncertainty propagation and fuzzy simulation described above is used to obtain the mean value and standard deviation of a financial efficient parameter for a given portfolio of project. To find an efficient project portfolio, the *PSM* model is used. To solve this model, a lexicographic approach is used. As a result, a set of Pareto-optimal solutions is obtained.

4. NUMERICAL EXAMPLE

The capital budget was determined for the production process presented in Fig. 1. This setup includes the production cycle in a steel industry, starting from the pig iron production, through the production of steel, hot rolling products to the production of products coated with metal and plastic.

Five investment projects are taken into consideration: steel making plant, hot rolled sheet mill, cold-rolled sheet mill, hot-dip galvanizing sheet plant and sheet organic coating plant. In Fig. 1 they are denoted with the suffix "-project". The gross profit (12) is used as a measure of efficiency of a portfolio of investment projects.

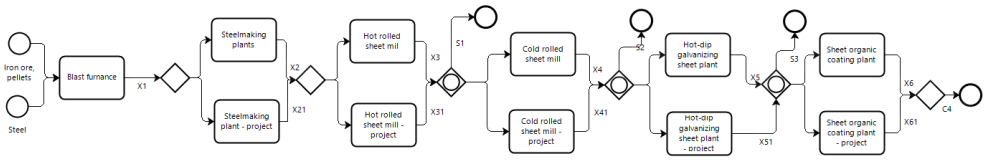

Fig. 1. Diagram of the considered technological setup

Decision variables for the estimation of the efficiency and risk of investment projects in the case of the iron metallurgy are the following: quantity and selling prices, costs of materials and quantity of investment outlays. The model parameters such as quantity of sale for each of products ranges being produced by the company, prices of these products, prices of metallurgic raw materials (prices of iron ores and the pellets), consumption per unit indexes and quantity of investment outlays are assumed to be uncertain, which makes the model closer to reality. The remaining parameters of the efficiency calculus are assumed to be deterministic. Moreover, it is assumed prices of individual assortment of metallurgical products and metallurgical raw materials are correlated. The same concerns sale quantities of each assortment of metallurgical products. These dependencies are taken into account when processing the values of uncertain parameters of the efficiency calculus.

Table 1. Trapezoidal fuzzy numbers (TFN) indicating material consumption

Material consumption	TFN
steel half-products – molten iron	(0.855; 0.860; 0.870; 0.875)
half-products – hot rolled steel sheets	(1.058; 1.064; 1.075; 1.078)
hot rolled steel sheets – cold rolled sheets	(1.105; 1.111; 1.124; 1.130)
cold rolled sheets – dip galvanized sheets	(1.010; 1.020; 1.026; 1.031)
dip galvanized sheets – organic coated sheets	(0.998; 0.999; 1.000; 1.001)

Table 2. Trapezoidal fuzzy numbers (TFN) for prices

Price	TFN (USD/t)
Iron ore	(335; 360; 400; 425)
Lumps	(375; 400; 440; 470)
Steel crap	(940; 960; 1010; 1035)
Hot rolled sheet	(2040; 2080.8; 2177.7; 2228.7)
Cold rolled sheet	(2220.08; 2266.65; 2370.15; 2427.08)
hot dip galvanized sheets and strips	(2535.75; 2588.25; 2709; 2772)
organic coated sheets and tapes	(3450.6; 3519.82; 3684.9; 3754.13)

Table 3. Probability distributions indicating sale parameters

Sale	Mean value	Std. dev.
hot rolled sheets	4704.0	117.5
cold rolled sheets	2750.0	51.4
hot dip galvanized – sheets and tapes	1147.9	52.4
organic coated – sheets and tapes	708.4	30.8

Material consumption as well as product and half-product prices are given in the form of fuzzy numbers. They are presented, respectively, in Table 1 and Table 2. Sale parameters are described by normal probability distributions given in Table 3. The Cholesky matrix which is used to process dependencies between sale parameters is the following:

$$\begin{pmatrix} 1.00000 & 0.87786 & 0.91142 & 0.86321 \\ 0.00000 & 0.47891 & 0.24007 & 0.27276 \\ 0.00000 & 0.00000 & 0.33418 & 0.34165 \\ 0.00000 & 0.00000 & 0.00000 & 0.25249 \end{pmatrix}.$$

In this computational example, the number of α-levels of fuzzy variables is set at 10 and the number of simulation is set to 100. The results obtained are presented in Fig. 2.

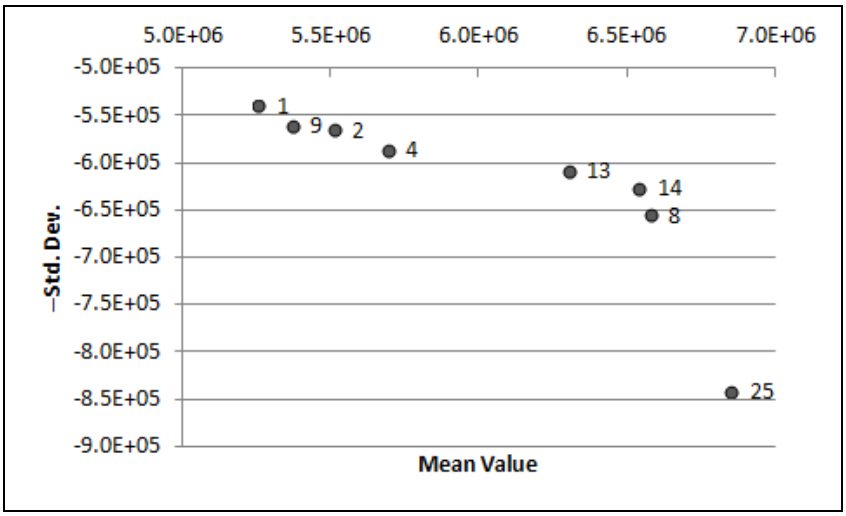

Fig. 2. Pareto optimal solutions for the capital budgeting problem

5. CONCLUSION

This paper presents a new method for selecting an effective portfolio of investment project. The presented concept of the mathematical model and numerical algorithm make it possible to generate a set of Pareto optimal solutions. These solutions are different variations of a company's acceptable capital budgets along with the estimation of their effectiveness (expected value of financial evaluation parameter) and risk (standard deviation of financial evaluation parameter). The method allows to take into account statistical as well as economic dependencies between projects. It also allows for flexible definition of uncertainty of the parameters using probability distribution or fuzzy numbers.

REFERENCES

[1] BAUDRIT C., DUBOIS D., GUYONET D., *Joint Propagation and Exploitation of Probabilistic and Possibilistic information in Risk Assessment*, IEEE Transaction on Fuzzy Systems, vol. 14, 2006, pp. 593–607. (DOI:10.1109/TFUZZ.2006.876720)

[2] CHAN F.T.S.., CHAN M.H., TANG N.K.H.: *Evaluation methodologies for technology selection*, Journal of Materials Processing Technology, 107, 2000: pp. 330–337

[3] FERSON S., GINZBURG L. R., *Difference method are needed to propagate ignorance and variability*", Reliab. Eng. Syst. Safe, vol. 54, 1996, pp. 133–144. (DOI:10.1016/S0951-8320(96)00071-3)

[4] GUYONNET D., BOURGINE B., DUBOIS D., FARGIER H., CME B., CHILS P. J., *Hybrid Approach for addressing uncertainty in risk assessment*, Journal of Environmental Engineering, vol. 126, 2003, pp. 68–76. (DOI:10.1061/(ASCE)0733-9372(2003)129:1(68))

[5] HUANG X., *Chance-constrained programming models for capital budgeting with NPV as fuzzy parameters*, Journal of Computational and Applied Mathematics, 198(1), 2007, pp. 149–159. doi:10.1016/j.cam.2005.11.026

[6] LIU Y-K, LIU B. *Fuzzy random variables: A scalar Expected Value Operator*. *Fuzzy Optimi-zation and Decision Making* 2, pp. 143-160, 2003.

[7] RĘBIASZ, B.: *Selection of efficient portfolios-probabilistic and fuzzy approach, comparative study*. Computers & Industrial Engineering, 2013; 64: pp. 1019–1032

[8] RĘBIASZ B., GAWEŁ B., SKALNA I.: *Fuzzy multiattribute evaluation of investments*, Computer Science and Information Systems (FedCSIS), 2013 Federated Conference on. IEEE, 2013.

[9] SEITZ N.E.: *Capital Budgeting and Long-Term Financing Decisions*, Third Edition, Dryden Press, USA, 1999

[10] YANG I.-T., *Simulation-based estimation for correlated cost elements*, International Journal of Project Management, vol. 23(4), 2005, pp. 275–282. (DOI:10.1016/j.ijproman.2004.12.002)

[11] ZULUAGA, A., SEFAIR J., MEDAGLIA A.: *Model for the Selection and Scheduling of Interdependent Projects*. Proceedings of the 2007 Systems and Information Engineering Design Symposium, University of Virginia., 2007.

Industry attractiveness, competitiveness, region attractiveness, method of principal components, eigenvalues, eigenvectors

Adriana HALIKOWSKA*
Daniel HALIKOWSKI**

ATTRACTIVENESS STUDY OF SPECIFIC INDUSTRY IN OPOLE VOIVODSHIP AND POLAND

In times of ubiquitous globalization every well-functioning company needs to know how and where to invest its capital. Often one of the ways of investing is opening new branches, that are supposed to become an important pillar of the whole enterprise. To achieve this goal the company must know which regions will bring profit, and which will not. One of the factors that affect the success of investment is industry attractiveness in the region. The authors used data published by Central Statistical Office of Poland to determine relation which allows to answer to the question whether chosen by the company industry is competitive. For the calculations authors used principal component method based on eigenvectors and eigenvalues. On this method basis the degree of explanation of attractiveness for Poland, Opole Voivodship and for the concentration coefficient was determined.

1INTRODUCTION

Competing is an inherent attribute of the market accompanying each company. However, before the company begins to operate on the market, the region should be examined. It should be examined whether there is a need for our products and services. Often companies, by improper placement of their business, lost liquidity what in turn led to bankruptcy. Therefore, proper market research is important, especially industry attractiveness in a particular region. On this basis it is much easier to decide whether to invest in new branch opening in the area.

* The University of Applied Sciences in Nysa, Institute of Finance, e-mail: adriana.halikowska@pwsz.nysa.pl

** The University of Applied Sciences in Nysa, Institute of Computer Science, e-mail: daniel.halikowski@pwsz.nysa.pl

1. INDUSTRY COMPETITIVENESS VERSUS INDUSTRY ATTRACTIVENESS

Competitiveness is a feature which implies relation of examined object to other objects. Thus, it has a comparative property. An interesting approach to classification of factors of system competitiveness presents Meyer-Stamer [7] (Fig. 1.). He described competitiveness as ability of regional level to generate high and rising incomes and growth of livelihood costs of its inhabitants.

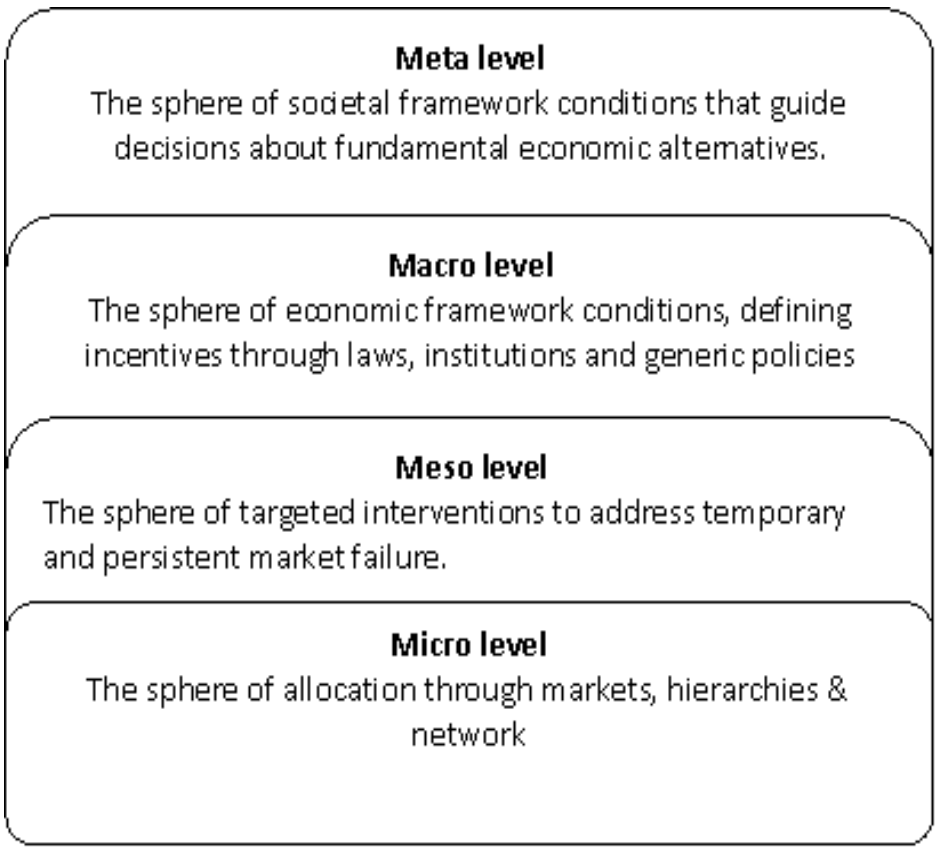

Fig. 1. Four Levels of Systemic Competitiveness
Source: MEYER-STAMER J., Systemic Competitiveness and Local Economic Development, [in:] Large Scale Systemic Change: Theories, Modelling and Practices, S. Bodhanya, eds., Duisburg, 2008.

The level at which the authors investigated the issue is the meso level. It is associated with industries and sectors. Therefore, these are fragments of the national economy, in this case Poland.

Criteria of division into subgroups is industry and territory. Gorynia and Łaźniewska [2, s.155] write that usage of competitiveness concept for provinces is less

adequate. Taking into consideration, the represented by the region, conditions of companies location, the term of competitiveness must be replaced by the term of attractiveness. Authors of this paper decided to change names, because of the necessity of examining of a collection of traits that determine the attractiveness of the industry in a given area. For analysis data from http://stat.gov.pl/ [10] and for the period of 2005-2012 were used.

An analysis of data relating to the industry "Wholesale and retail trade; repair of motor vehicles and motorcycles" was made. The choice of sector largely depended on the availability of the data. The collected data were used to analyze the attractiveness of Opole Voivodship, Poland and the concentration coefficient at three dimensions. In the process of searching for indicators that have been used to study industry, the problem was divided into several sub-areas. Among other things, they were economic operators, market, people, innovation. As a result of extensive research eight indicators has been identified. The task of these indicators is to best describe the attractiveness of the examined industry.

1.1. INDUSTRY ATTRACTIVENESS FOR POLAND

The first part of the research was the study of the attractiveness of the industry "Wholesale and retail trade; repair of motor vehicles and motorcycles" for Poland. The study used the method of eigenvalues and eigenvectors. This method basing on eigenvectors allows to determine which of the eigenvalues explains the investigated phenomenon to the greatest extent. Moreover, this method allows to specify the explanation of the phenomenon in the percentage.

For Poland, after insertion of the collected data, the following matrix of eigenvalues was achieved:

$$\begin{bmatrix} 0{,}374 & -0{,}395 & -0{,}181 & -0{,}031 & 0{,}758 & -0{,}233 & 0{,}198 & -0{,}054 \\ 0{,}392 & 0{,}320 & 0{,}421 & -0{,}412 & -0{,}341 & 0{,}261 & 0{,}536 & 0{,}163 \\ -0{,}099 & 0{,}462 & 0{,}662 & 0{,}561 & 0{,}114 & -0{,}069 & 0{,}068 & -0{,}017 \\ -0{,}144 & -0{,}718 & 0{,}075 & 0{,}526 & -0{,}340 & 0{,}033 & 0{,}225 & 0{,}121 \\ -0{,}377 & 0{,}244 & -0{,}388 & 0{,}048 & 0{,}172 & -0{,}177 & 0{,}744 & 0{,}176 \\ 0{,}433 & 0{,}116 & -0{,}192 & 0{,}179 & -0{,}338 & -0{,}495 & 0{,}168 & -0{,}585 \\ 0{,}443 & 0{,}1396 & -0{,}176 & 0{,}190 & -0{,}097 & -0{,}338 & -0{,}168 & 0{,}758 \\ 0{,}399 & 0{,}148 & -0{,}358 & 0{,}408 & 0{,}166 & 0{,}693 & 0{,}089 & -0{,}090 \end{bmatrix} \quad (1)$$

For the eigenvectors, which have been presented (1), a vector of eigenvalues that determine the extent to which the individual vectors of the matrix (2) explain the attractiveness in the test industry, can be created. Eigenvalues can be represented in the following form:

$$[4{,}9402 \quad 1{,}699993 \quad 1{,}30522 \quad 0{,}0498 \quad 0{,}0048 \quad 1{,}54E-0{,}5 \quad 7{,}49E-0{,}7 \quad 3{,}47E-17] \quad (2)$$

To determine the percentage extent to which individual vectors describe the industry attractiveness, their values were compared to their sum.

The percentage record of vector of eigenvalues is as follows:

$$[0{,}62 \quad 0{,}21 \quad 0{,}16 \quad 0{,}0062 \quad 0{,}000599 \quad 1{,}93E-0{,}6 \quad 9{,}35E-0{,}8 \quad 4{,}3433E-18] \quad (3)$$

Equation 3 shows that the most important are the first three eigenvalues. They describe the phenomenon in 98%. Other eigenvectors explain the phenomenon in a minor degree. It is therefore necessary to reduce the vector of eigenvalues and write it in the following form:

$$Eigenvalues = [62\% \quad 21\% \quad 16\%] \quad (4)$$

Similarly it must be proceeded with eigenvectors by cutting off the last 5 columns. In the final record the eigenvectors matrix looks as follows:

$$\begin{bmatrix} 0{,}374 & -0{,}395 & -0{,}181 \\ 0{,}392 & 0{,}320 & 0{,}421 \\ -0{,}099 & 0{,}462 & 0{,}662 \\ -0{,}144 & -0{,}718 & 0{,}075 \\ -0{,}377 & 0{,}244 & -0{,}388 \\ 0{,}433 & 0{,}116 & -0{,}192 \\ 0{,}443 & 0{,}1396 & -0{,}176 \\ 0{,}399 & 0{,}148 & -0{,}358 \end{bmatrix} \quad (5)$$

Then, breaking up the whole matrix into individual equations, we obtain the following equation that explains the industry attractiveness in approximately 62%:

$$Y_1 = 0{,}374 \bullet X_1 + 0{,}392 \bullet X_2 - 0{,}099 \bullet X_3 - 0{,}144 \bullet X_4 - 0{,}377 \bullet X_5 + 0{,}433 \bullet X_6 + 0{,}443 \bullet X_7 + 0{,}399 \bullet X_8 \quad (6)$$

For the assumption that the industry attractiveness is explained in approx. 21% the following equation will be obtained:

$$Y_2 = -0{,}395 \bullet X_1 + 0{,}320 \bullet X_2 + 0{,}462 \bullet X_3 - 0{,}718 \bullet X_4 + 0{,}244 \bullet X_5 + 0{,}116 \bullet X_6 + 0{,}1396 \bullet X_7 + 0{,}148 \bullet X_8 \quad (7)$$

When the industry attractiveness is explained in 16% the equation will look like this:

$$Y_3=-0{,}181 \bullet X_1+0{,}421 \bullet X_2+0{,}662 \bullet X_3+0{,}075 \bullet X_4-0{,}388 \bullet X_5-0{,}192* \bullet X_6-0{,}176 \bullet X_7-0{,}358 \bullet X_8 \quad (8)$$

Combining equations (6), (7) and (8) into one we get formula for calculating the attractiveness of the industry "Wholesale and retail trade; repair of motor vehicles and motorcycles" for Poland in the form of a matrix, and it looks as follows:

$$[Y_1 \;\; Y_2 \;\; Y_3] = [X_1 \;\; X_2 \;\; X_3 \;\; X_4 \;\; X_5 \;\; X_6 \;\; X_7 \;\; X_8] \bullet \begin{bmatrix} 0{,}374 & -0{,}395 & -0{,}181 \\ 0{,}392 & 0{,}320 & 0{,}421 \\ -0{,}099 & 0{,}462 & 0{,}662 \\ -0{,}144 & -0{,}718 & 0{,}075 \\ -0{,}377 & 0{,}244 & -0{,}388 \\ 0{,}433 & 0{,}116 & -0{,}192 \\ 0{,}443 & 0{,}1396 & -0{,}176 \\ 0{,}399 & 0{,}148 & -0{,}358 \end{bmatrix} \quad (9)$$

Explanation of the model:
X_1 - gross profit
X_2 - expenditure on fixed assets
X_3 - the reference rate
X_4 - the number of people of working age
X_5 - the number of unemployed
X_6 - the average monthly expenditure on 1 person
X_7 – the registered sales in the automotive sector
X_8 - higher education alumni per 10 thousand population

The model presented in the form of matrix [1.1] clearly shows that the variables: gross profit (X_1), expenditure on fixed assets (X_2), the average monthly expenditure on 1 person (X_6), the registered sales in the automotive sector (X_7) and higher education alumni per 10 thousand population (X_8) have a positive impact on the dependent variable Y_1. Therefore, they are attractiveness stimulants. The greatest stimulant of these variables is a variable X_7 = 0.443, ie. sales in the automotive sector. Among destimulants that have a negative impact on the dependent variable Y_1 variables X_3, X_4 and X_5 should be listed. The most significant negative impact has X_5 = -0.377 (the number of unemployed). This situation occurs for the equation where the attractiveness is explained in 62%, that is the eigenvalue is 62%.

When we look at the situation when the dependent variable Y_2 has been explained in 21% the differences can be seen. In this case the greatest stimulant is explanatory variable X_3 (the reference rate), while the largest destimulant is X_4 (the number of people of working age). In case of X_4 it can be seen that this influence is very big, because increase of attractiveness by the unit causes decrease in number of people of working age about 0,718.

Among the analyzed dependent variables there is also Y_3, which is an explaining system of the dependent variable in 16%. In this arrangement, the greatest positive impact on the dependent variable has the reference rate (X_3) amounting to 0,662, while the largest negative impact has the number of unemployed persons (X_5) and amounts to -0,388.

1.2 ATTRACTIVENESS FOR OPOLE VOIVODSHIP

The second part of the research was the study of attractiveness sector "Wholesale and retail trade; repair of motor vehicles and motorcycles" for Opole Voivodship. As in the case of Poland the method of eigenvectors and eigenvalues was used. The results were analyzed and it was found that the attractiveness for Opole Voivodship is explained in more than 90%.

For Opole Voivodship, after insertion of the collected data, the following matrix of eigenvalues was achieved:

$$\begin{matrix} 0{,}350 & 0{,}282 & 0{,}484 & 0{,}468 & 0{,}539 & -0{,}175 & -0{,}133 & 0{,}066 \\ 0{,}398 & 0{,}108 & -0{,}142 & -0{,}672 & 0{,}227 & -0{,}539 & -0{,}037 & -0{,}120 \\ -0{,}159 & 0{,}678 & -0{,}660 & 0{,}194 & 0{,}168 & -0{,}007 & 0{,}044 & 0{,}109 \\ -0{,}379 & 0{,}197 & 0{,}332 & -0{,}460 & 0{,}271 & 0{,}217 & 0{,}115 & 0{,}601 \\ -0{,}318 & -0{,}506 & -0{,}308 & 0{,}042 & 0{,}601 & -0{,}023 & -0{,}429 & -0{,}012 \\ 0{,}402 & 0{,}072 & -0{,}082 & -0{,}225 & 0{,}258 & 0{,}789 & -0{,}020 & -0{,}294 \\ 0{,}401 & -0{,}033 & -0{,}150 & 0{,}015 & -0{,}312 & 0{,}091 & -0{,}608 & 0{,}583 \\ 0{,}354 & -0{,}385 & -0{,}274 & 0{,}163 & 0{,}178 & -0{,}010 & 0{,}641 & 0{,}426 \end{matrix} \quad (10)$$

In this case eigenvalues can be represented in the following form:

$$[0{,}76 \quad 0{,}16 \quad 0{,}078902 \quad 0{,}000524 \quad 0{,}000346 \quad 5{,}27E-06 \quad 3{,}76E-09 \quad 1{,}56E-17] \quad (11)$$

For Opole Voivodship the first vector is explained in 76%. Thus, it determines the sector attractiveness more accurately than the first vector created for Poland. The general form of the formula for Y'_1 takes the following form:

$$Y'_1 = 0{,}350 \bullet X_1 + 0{,}398 \bullet X_2 - 0{,}159 \bullet X_3 - 0{,}379 \bullet X_4 - 0{,}319 \bullet X_5 + 0{,}402 \bullet X_6 + 0{,}401 \bullet X_7 + 0{,}354 \bullet X_8 \quad (12)$$

While, in the second vector, which describes sector attractiveness for Opole Voivodship the level of explanation is 16%.

$$Y'_2 = 0{,}282 \bullet X_1 + 0{,}108 \bullet X_2 + 0{,}678 \bullet X_3 + 0{,}197 \bullet X_4 - 0{,}506 \bullet X_5 + 0{,}072 \bullet X_6 - 0{,}033 \bullet X_7 - 0{,}385 \bullet X_8 \quad (13)$$

In case for Poland and for Opole Voivodship the attractiveness can be presented as one equation, that is based on equations (12) and (13):

$$[Y'_1 \quad Y'_2] = [X_1 \quad X_2 \quad X_3 \quad X_4 \quad X_5 \quad X_6 \quad X_7 \quad X_8] \bullet \begin{bmatrix} 0{,}350 & 0{,}282 \\ 0{,}398 & 0{,}108 \\ -0{,}159 & 0{,}678 \\ -0{,}379 & 0{,}197 \\ -0{,}319 & -0{,}506 \\ 0{,}402 & 0{,}072 \\ 0{,}401 & -0{,}033 \\ 0{,}354 & -0{,}385 \end{bmatrix} \quad (14)$$

Assuming that the eigenvalues of the matrix (14) are as follows:

$$Eigenvalues = [76\% \quad 16\%] \quad (15)$$

In this case the variable Y'_1 is explained in 76%, whereas Y'_2 in 16%. Carrying out a similar analysis (like in the case of Poland), it can be seen that in Opole Voivodship an increase in the explanatory variable Y'_1 by the unit causes an increase in the gross profit (X_1), expenditures on fixed assets (X_2), the average monthly expenditure on 1 person (X_6), sales in the automotive sector (X_7) and the higher education alumni per 10 thousand population (X_8). Conversely to the explanatory variable are changes of reference rate (X_3), the number of people of working age (X_4) and the number of unemployed (X_5). Among the explanatory variables explaining the phenomenon of attractiveness in the sector of "Wholesale and retail trade; repair of motor vehicles and motorcycles" for Opole Voivodship the greatest stimulant for Y'_1 are average monthly expenditure on 1 person (X_6 = 0,402) and registered sales in the automotive sector (X_7 = 0,401). While, the largest destimulant in the examined case is the number of people of working age (X_4 = 0,379).

1.3 ATTRACTIVENESS FOR CONCENTRATION COEFFICIENT

The last stage of the study was the concentration coefficient investigation. The whole procedure was similar like in the case for Poland and for Opole Voivodship. However, as input data the concentration coefficient was used. The concentration coefficient task is to explore the degree to which Opole Voivodship is responsible for participating in the attractiveness of the tested industry in Poland.

In the basic version the equation presents as follows:

$$\begin{matrix} 0{,}272 & 0{,}398 & 0{,}705 & 0{,}113 & 0{,}239 & 0{,}343 & -0{,}281 & 0{,}069 \\ 0{,}434 & 0{,}124 & -0{,}049 & -0{,}057 & -0{,}049 & -0{,}312 & -0{,}169 & -0{,}814 \\ -0{,}154 & 0{,}554 & -0{,}557 & 0{,}217 & 0{,}463 & -0{,}034 & -0{,}303 & 0{,}069 \\ -0{,}435 & 0{,}063 & 0{,}217 & -0{,}191 & -0{,}337 & -0{,}425 & -0{,}647 & 0{,}095 \\ -0{,}337 & -0{,}402 & 0{,}276 & 0{,}027 & 0{,}727 & -0{,}272 & 0{,}042 & -0{,}208 \\ -0{,}244 & 0{,}530 & 0{,}259 & 0{,}186 & -0{,}129 & -0{,}432 & 0{,}595 & -0{,}028 \\ -0{,}433 & -0{,}102 & 0{,}024 & 0{,}637 & -0{,}256 & 0{,}394 & -0{,}085 & -0{,}410 \\ 0{,}404 & -0{,}249 & 0{,}013 & 0{,}678 & -0{,}040 & -0{,}433 & -0{,}138 & 0{,}327 \end{matrix} \quad (16)$$

The percentage record of vector of eigenvalues is as follows:

$$[0{,}64 \quad 0{,}29 \quad 0{,}06 \quad 0{,}005878 \quad 0{,}000432 \quad 8{,}8E-05 \quad 5{,}85E-05 \quad 1{,}27E-17] \quad (17)$$

Also this time the result of the study were two explained variables Y''_1 and Y''_2. In general, the phenomenon has been described in 85%. Other eigenvectors are less significant.

For the first variable the degree of explanation of the phenomenon of the attractiveness is 69%. It can be described by the following equation:

$$Y''_1 = 0{,}273*X_1+0{,}434*X_2-0{,}154*X_3-0{,}435*X_4-0{,}337*X_5-0{,}244*X_6-0{,}433*X_7+0{,}404*X_8 \quad (18)$$

If the degree of explanation of attractiveness for concentration coefficient is about 16% the equation will be as follows:

$$Y''_2 = 0{,}399*X_1+0{,}124*X_2+0{,}554*X_3+0{,}063*X_4-0{,}402*X_5+0{,}530*X_6-0{,}102*X_7-0{,}249*X_8 \quad (19)$$

Analyzing approach, where we examine the extent to which Opole Voivodship is responsible for participating in the attractiveness of the tested industry on the background of Poland, patterns (18) and (19) can be expressed in the following form:

$$[Y''_1 \quad Y''_2] = [X_1 \quad X_2 \quad X_3 \quad X_4 \quad X_5 \quad X_6 \quad X_7 \quad X_8] \bullet \begin{bmatrix} 0{,}273 & 0{,}399 \\ 0{,}434 & 0{,}124 \\ -0{,}154 & 0{,}554 \\ -0{,}435 & 0{,}063 \\ -0{,}337 & -0{,}402 \\ -0{,}244 & 0{,}530 \\ -0{,}433 & -0{,}102 \\ 0{,}404 & -0{,}249 \end{bmatrix} \quad (20)$$

In this case, the vector of eigenvalues, which determine the degree of attractiveness explanation is as follows:

$$Eigenvalues = [69\% \quad 16\%] \tag{21}$$

Dependent variable $Y"_1$ is explained in approx. 69%. In the approach, where the concentration ratio is examined, it can be clearly seen that the explanatory variables are mostly negatively correlated with the response variable (5 out of 8 variables). Two explanatory variables are destimulants that have maximum impact on the dependent variable, that is X_4 (number of people of working age) and X_7 (the registered sales in the automotive sector). With the increase of $Y"_1$ by the unit the mentioned explanatory variables decrease by approx.0,4. Among the positively correlated variables the greatest positive impact have variables: X_2 (expenditure on fixed assets) and X_8 (higher education alumni per 10 thousand population). Their value fluctuates around 0,4.

When the explanatory variable is explained in 16% ($Y"_2$) the situation is slightly different. To the greatest extent the explanatory variable is explained by the reference rate (X_3), which assumes a value of 0,554 and is positively correlated with the explanatory variable. Besides, X_6 (average monthly expenses on one person) also should be mentioned. X_6 together with an increase of the explanatory variable by one unit changes by 0,530. While the biggest destimulant in this case is the number of unemployed (X_5).

2. SUMMARY

Every company that wants to develop needs to make various investments. One of these is certainly opening of new branches. Such activities are very risky. In order to minimize the risk the analysis of the competitiveness of the industry, in which we want to invest in a particular region, can be done. The proposed in the article methodology allows to do this. The analysis of the data shows that the attractiveness is influenced by various factors, and knowing their distribution, it is possible to assess with a certain probability their behavior in the nearest future.

The proposed by the authors approach is universal. It can be applied to any industry and for every voivodship. The only limitation that may occur is lack of data. However, in today's era of computerization most of the data is stored in an electronic version, and year after year the database is becoming more extensive.

REFERENCES

[1] DĘBSKI M., *Kreowanie silnej marki*., Warszawa, PWE, 2009.
[2] Red. nauk. GORYNIA M., ŁAŹNIEWSKA E., *Kompendium wiedzy o konkurencyjności*, Warszawa, PWN, 2009. http://zif.wzr.pl/pim/2012_4_2_2.pdf
[3] KALETA A., MOSZKOWICZ K., *Zarządzanie strategiczne w badaniach teoretycznych i w praktyce*, Wydawnictwo: Uniwersytet Ekonomiczny We Wrocławiu, 2008.
[4] KAPLAN R.S., NORTON D. P., *Wdrażanie strategii dla osiągnięcia przewagi konkurencyjnej,* Warszawa, PWN, 2010.
[5] KIM D., MARION B.W., *Domestic Market Structure and Performance in Global markets: Theory and Empirical Evoidence from U.S. Food Manufacturing Industries*, Review of Industrial Organization, Vol. 12, 1997, pp.335–354. http://www.aae.wisc.edu/fsrg/publications/Archived/wp-109.pdf
[6] MATYJAS Z., *Wpływ czynników oddziałujących na poziomie firmy oraz czynników sektorowych na wyniki finansowe spółek w świetle badań empirycznych*, 2012.
[7] MEYER-STAMER J., *Systemic Competitiveness and Local Economic Development, [in:] Large Scale Systemic Change: Theories, Modelling and Practices*, S. Bodhanya, eds., Duisburg, 2008.
[8] PORTER MICHAEL E., *Strategia konkurencji.*, Warszawa, PWE, 1992.
[9] STANKIEWICZ M. J., *Konkurencyjność przedsiębiorstw- nowe podejście*, Wydawnictwo Naukowe PWN, Warszawa- Poznań, 2002.
[10] http://stat.gov.pl/

Merges & acquisitions, goodwill
investment evaluation
value based management

Michal J. KOWALSKI*

VALUE CREATION MEASURES AS A SUPPORT IN INVESTMENT DECISIONS CONCERNING MERGES AND ACQUISITIONS

The aim of the article is to evaluate the possibility of value measures (residual income, EVA) use as a support in investment decisions concerning merges and acquisitions. A proposal of operating procedure in acquisition decision-making was presented. A special attention was paid to transactions where only part of shares are acquired. A detailed propositions of value measures adjustments concerning acquisition accounting were also included. A detailed solutions for goodwill, badwill and restructuring costs were proposed. An example of the proposed approach was given.

1. INTRODUCTION

1.1. VALUE CREATION IN MERGERS AND ACQUISITIONS

Mergers and acquisitions are, beside the organic growth, one of the main paths to increase capital value. They are currently presented as a natural part of corporate practice, and as such appear in strategy textbooks and MBA classes [6]. Numerous studies conducted on the processes of mergers and acquisitions show that their effectiveness is not always satisfactory and, as a result, there is still a need to further explore methods supporting investment decision-making in this area.

Business acquisition does not differ significantly from other kinds of investments. From the investor's point of view, it is a cash-out and cash-in process; an additional aspect is the target company valuation.

Income approach can be modified by its use not on the cash flows but on the value measures, especially Economic Value Added (EVA). As it has been shown on numer-

* Wroclaw University of Technology, Institute of Organization and Management

ous occasions, both approaches lead to identical results and generally the same conclusions [2, CH 6-6]. However, as demonstrated by Dilon & Owers use of EVA as the basis for NPV calculation allows for rewarding divisional performance [4]. EVA measures identifies a degree to which a company is successful in earning rates of return that exceed its cost of capital [**Błąd! Nie można odnaleźć źródła odwołania.**]. EVA valuations provide insights on value enhancement because of its focus on excess returns defined in terms of return and cost of capital [**Błąd! Nie można odnaleźć źródła odwołania.**]

While using value measures it is of great importance to remember about data adjustments in financial reporting. The purpose of these adjustments is to modify the accounting data in order to ensure the best representation of the business and economic sense, which can be distorted by some procedures resulting from restrictive accounting law. The adjustments have been widely presented in the literature, where the interested reader is sent[11]. It should be noted that the adjustments ought not to change the value in the whole period of its application but aim to present economic sense. Therefore, it can be concluded that the NPV of the adjustments, taking into account the full period of their use, amounts to zero [5]. Only under this condition, the indicated above regularity that NPV and DCF used in relation to FCF and EVA produce the same results.

One group of adjustments made in the calculation of EVA constitutes goodwill adjustments [11]. However, they only concern write-offs of goodwill [10], and do not relate to acquisition transactions and the goodwill impact on investment evaluation in merges and acquisitions. Acquisition transaction settlement is regulated, in detail, in the international accounting law IAS 36, IAS 38, IFRS 3 and IFRS 13. Separate procedures are used for acquisition of goodwill and badwill [3].

The use of value creation measures is usually associated with their subsequent use in incentive systems, which constitute one of the elements of VBM (Value-Based Management) systems [9]. In this context, the search for presentation methods of financial results, in accordance with the economic sense, i.e. introducing appropriate accounting adjustments, is of particular importance. If these results are the direct basis for remuneration, they must on the one hand encourage managers to act, and on the other to provide an effective increase in the value for the investor making the mergers and acquisitions transaction.

2. VALUE MEASURES FOR MERGERS AND ACQUISITIONS

2.1. PROPOSAL OF OPERATING PROCEDURE

Figure 1 presents a procedure for target company valuation for the bidder using measures of value creation.

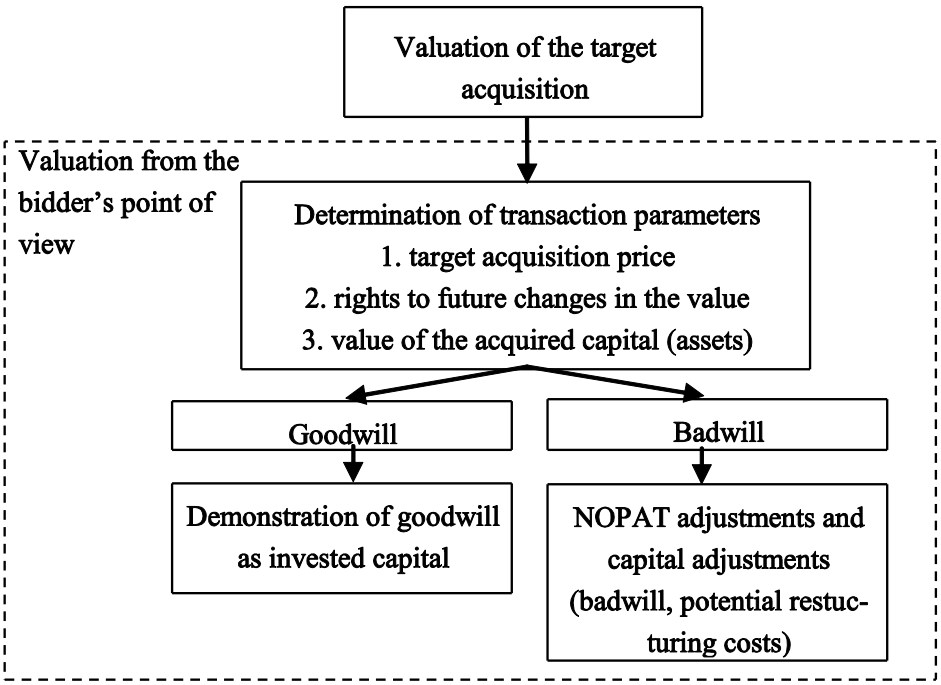

Fig. 1. Stages of mergers and acquisitions using value measures

The valuation will take place in two stages. In the first stage, the target company will be evaluated, as well as the future cash flows or changes in value. This stage is to determine the equity value of the target company.

In the second stage, the evaluation will be made from the bidder's point of view. In the case of application of value creation measures, the valuation will be made by comparing the invested capital and future changes in value, subjected to discounting. At this stage, it is important to determine: (1) what % of the value generated by the target company (EVA^A) will belong to the investor, and (2) the amount of the invested capital by the bidder. The answer for the first question comes down to the investor's share in the acquired equity. Given that the measures of NPV are additive [7][1], an investor who has not acquired 100% of shares owns appropriate part of operating profit after tax of the target company ($NOPAT^A$) and appropriate part of invested capital (IC^A), because

$$\%EVA^A = \%NOPAT^A - \%IC^A \bullet WACC \tag{1}$$

[1] Under the condition of constant WACC in the analysed period

where:

% – investor's shares in the acquired equity

EVA^A – Economic Value Added of the acquired company

$NOPAT^A$ – Net Operating Profit After Tax of the acquired company

While answering the second question, it must be noted that from the bidder's point of view its invested capital cannot be reduced only to the resulting ratio of the acquired shares of the invested capital (% IC^A). The investment will be equal to the expenditure (price) on the acquisition of shares (P).

At this point, an important problem from the point of view of application of value measures arises, a problem related to the relationship between the expenditure on the purchase (price) of the shares, or parts of shares, and the capital the investor receives in return (assets of the company). There are three possible situations:

(1) P > %x IC

(2) P < %x IC

(3) P = %x IC

From the point of view of accounting law in case (1) we are dealing with goodwill, and in (2) badwill. Both situations require a separate analysis, in order to ensure transaction evaluation using values measures that will fully reflect economic sense and ensure proper evaluation. Case (3) as a purely theoretical does not require further analysis.

2.2. GOODWILL IN VALUE EVALUATION

In the case of goodwill (GW), from the investor's point of view their invested capital equals the amount of money spent on purchasing stocks/shares. Therefore, economic profit belonging to the investor in *i* period equals:

$$EVA_i = \% \, NOPAT_i^A - P \bullet WACC \tag{2}$$

alternatively:

$$EVA_i = \%NOPAT_i^A - (\%IC_i^A + GW) \bullet WACC = EVA_i^A + GW \bullet WACC \tag{3}$$

Then the equation for investment evaluation is:

$$NPV = \sum_{i=1}^{\infty} \frac{EVA_i^A + GWxWACC}{(1+WACC)^i} \tag{4}$$

2.3. INCLUDING BADWILL AND RESTRUCTURING COSTS IN VALUE EVALUATION

International accounting law includes badwill in the financial results for a given period [3]. Such an approach would be sufficient from the point of view of value measures if the company acquired at a price lower than the value of its assets was treated as a realization of a single market opportunity. Then badwill (BW) increases only once the operating result and does not affect invested capital, and

$$EVA_1 = \%NOPAT_1^A + BW - \%IC_1^A \bullet WACC \tag{5}$$

For companies, however, bargain acquisitions are not generally realized and badwill recognition usually requires recognition of restructuring costs (CR) in subsequent periods. In this situation it is proposed to:

– during the restructuring period adjust the restructuring costs, which should be added to the invested capital
– during the restructuring period reduce invested capital by badwill and do not account for it in the NOPAT
– after the restructuring period settle badwill with activated restructuring costs

As a result, from the investor's point of view its economic profits from the acquired company:

– in the restructuring period $i \in (0, j)$

$$EVA_i = \%NOPAT_i^A + CR_i - (\%IC_i^A + \sum_0^i CR_i - BW) \bullet WACC \tag{6}$$

– in the year after restructuration $j+1$

$$EVA_{j+1} = \%NOPAT_{j+1}^A + (BW - \sum_{i=0}^{j} CR_i) - \%IC_{j+1}^A \bullet WACC \tag{7}$$

– in subsequent periods:

$$EVA_{j+2} = \%EVA_{j+2}^A \tag{8}$$

Indicated in equations (6) and (7) the badwill adjustment and restructuring costs realize the NPV = 0 condition, and as a result do not affect the evaluation of the investment but only on presentation of individual periods. Equation of investment evaluation in case of acquisition with badwill takes the form:

$$NPV = \sum_{i=0}^{j} \frac{EVA_i}{(1+WACC)^i} + \frac{EVA_{j+1}}{(1+WACC)^i} + \sum_{t=j+2}^{\infty} \frac{EVA_t}{(1+WACC)^t} \tag{9}$$

3. APPLICATION EXAMPLE

3.1. ASSUMPTIONS

In this chapter, I will present an example of the use of VBM value measures in capital investment. The subject of the acquisition are 49% shares of the Company X for the price of 13.5 million PLN. Company's invested capital was set at 20 million PLN. The cost of capital in the example was set at 10%.

3.2. INVESTMENT EVALUATION

The first step in the evaluation of the investment is valuation of the subject of the investment, i.e. the target company. The valuation was made based on the DCF method. The basis of the valuation was the forecast made by the Company. Table 1 shows the measurement conducted on the basis of DCF based on FCFF.

Table 1. FCF Valuation – 100% shares

Year of forcast	1	2	3	4	5	…
EBIT	3,7	4,7	4,5	4,2	3,1	3,1
TAX	-0,7	-0,9	-0,8	-0,8	-0,6	-0,6
NetCapEx	3,9	1,7	1,0	2,2	2,9	
dWCR	-0,4	-0,7	-0,1	-0,1	-0,1	
FCF	6,5	4,9	4,6	5,5	5,4	2,5
Discount Factor	0,91	0,83	0,75	0,68	0,62	6,21
PV FCF	5,9	4,0	3,4	3,8	3,3	15,8
sum of PVFCF	36,2					
Eqiuty (100% shares)	36,2					

Where:
NetCapEx – capital expenses reduced by depreciation
dWCR – delta working capital requiment

Table 2 shows the calculation of EVA for the company under valuation. An integral part of the calculation was to determine the invested capital. IC has been included from the beginning of the period. Calculation of initial balance of invested capital was made on the base of balance sheet of the company in accordance with the principles of the calculation of EVA. Elements that according to the forecast impact on FCF adjust the level of capital with the opposite sign, for example, the depreciation reduces IC and capital expenses increase IC.

Table 2. EVA Calculation

Year of forcast	1	2	3	4	5	…
EVA	1,0	2,2	2,1	2,0	1,3	1,6
NOPAT	3,0	3,8	3,6	3,4	2,5	2,5
EBIT	3,7	4,7	4,5	4,2	3,1	3,1
TAX	-0,7	-0,9	-0,8	-0,8	-0,6	-0,6
CoC	-2,0	-1,7	-1,5	-1,5	-1,2	-1,0
IC	20,0	16,5	15,5	14,6	12,5	9,6
WACC	10%	10%	10%	10%	10%	10%
IC initial balance	20,0	16,5	15,5	14,6	12,5	9,6
NetCapEx	-3,9	-1,7	-1,0	-2,2	-2,9	0,0
dWCR	0,4	0,7	0,1	0,1	0,1	0,0
IC closing balance	16,5	15,5	14,6	12,5	9,6	9,6

Table 3 presents the valuation of 100% shares of the company conducted on the basis of discounting EVA. Initial balance of invested capital is added to the present value of future EVA. EVA outside the period of the forecast is the basis for the calculation of perpetuity by the classical model of Gordon.

Table 3. EVA Valuation -100% shares

Year of forcast	0	1	2	3	4	5	…
EVA		1,0	2,2	2,1	2,0	1,3	1,6
Discount Factor		0,909	0,826	0,751	0,683	0,621	6,209
PV (EVA)		0,9	1,8	1,6	1,3	0,8	9,8
sum of PV(EVA)	16,2						
IC initial balance	20,0						
Eqiuty (100% shares)	36,2						

Both methods of valuation, FCF valuation and EVA valuation, produced identical results. However, EVA indicates whether the target company will provide in the subsequent periods profit greater than the cost of capital employed. The amount of initial balance does not affect the valuation results but affects EVA achieved in subsequent years.

The next step in the process is to evaluate if the invested capital on terms specifically defined in the assumptions is profitable for the bidder. The acquisition should be treated as an investment. Expenditure on the acquisition of 49% is the cash out, in return for which the bidder acquires the right up to 49% of future FCF of the target company and 49% EVA, which will be generated in the future. Table 4 presents the

evaluation of investment income in the company's shares in accordance with the concept of NPV.

Table 4. Investment evaluation PV of FCF

Year of forcast	0	1	2	3	4	5	…
FCF for purchaser (49%)	-13,5	3,2	2,4	2,2	2,7	2,6	1,2
Discount Factor	1,00	0,91	0,83	0,75	0,68	0,62	6,21
PV (FCF dla GI)	-13,5	2,9	2,0	1,7	1,8	1,6	7,7
NPV	4,2						

The bidder can evaluate the investment from the point of view of EVA, which the acquisition will generate, but assuming that the worked-out residual income must cover the entire invested capital by the investor at the time of purchase of shares. Therefore, it is necessary to compare the purchase price of assets with the invested capital, for an investor (purchaser). The bidder acquires 49% shares, therefore, 9.8 million capital employed. From the point of view of EVA goodwill will be 3.7 million (i.e. 13.5 - 9.8 = 3.7). Such an approach means that in the calculation of EVA for the investor invested capital presents the price spent on the acquisition.

Table 5. Investment evaluation PV of EVA with goodwill

Year of forcast	0	1	2	3	4	5	…
EVA for purchaser		0,1	0,7	0,6	0,6	0,3	0,4
NOPAT (49%)		1,5	1,9	1,8	1,7	1,2	1,2
CoC		-1,4	-1,2	-1,1	-1,1	-1,0	-0,8
IC		13,5	11,8	11,3	10,8	9,8	8,4
IC initial balance		13,5	11,8	11,3	10,8	9,8	8,4
IC 49%		9,8	8,1	7,6	7,1	6,1	4,7
Goodwill		3,7	3,7	3,7	3,7	3,7	3,7
Discount Factor	1,00	0,91	0,83	0,76	0,68	0,62	6,21
PV EVA	0,0	0,1	0,6	0,5	0,4	0,2	2,5
Sum of PV(EVA)	4,2						

Valuation of the investment using EVA led to identical results as the classical procedure i.e. NPV based on cash flows. Analysis from the point of view of EVA allows to assess the impact of the planned acquisition on the bidder's added value. The presented example shows that EVA achieved by the purchaser in the first year is close to zero. Thus, in the first year of acquisition, it has a little effect on the value of the buyer; it does not raise the value of the bidder's company.

3. 3. INVESTMENT EVALUATION IN THE CASE OF BADWILL

Assuming that the transaction was executed at a price lower than the 9.8 million PLN, i.e. 5.0 million PLN, we have to deal with the rise of badwill. The acquisition price is lower than the value of the property. For the readability of the example, let's assume that the transaction is possible under these conditions due to the need to restructure, which costs have been set at 1.5 million in three subsequent years. The accounting law requires recognition of badwill and restructuring costs as factors that reduce the income in the period. In this situation, the impact of investments on bidder's EVA and evaluation of the efficiency of investment are presented in Table 6.

Table 6. Investment evaluation PV of EVA with badwill and cost of restructuring

Year of forcast	0	1	2	3	4	5	…
EVA for purchaser		4,3	-0,4	-0,5	1,0	0,6	0,8
NOPAT (49%)		1,5	1,9	1,8	1,7	1,2	1,2
Cost of restructuring		-1,5	-1,5	-1,5			
Badwill		4,8					
Cost of Capital		-0,5	-0,8	-0,8	-0,7	-0,6	-0,5
IC		5,0	8,1	7,6	7,1	6,1	4,7
IC		5,0	8,1	7,6	7,1	6,1	4,7
IC 49%		9,8	8,1	7,6	7,1	6,1	4,7
Goodwill		-4,8	0,0	0,0	0,0	0,0	0,0
Discount Factor	1,00	0,91	0,83	0,75	0,68	0,62	6,21
PV EVA	0,0	3,9	-0,4	-0,4	0,7	0,4	4,8
Sum of PV(EVA) = NPV	9,0						

The estimated EVA of the investment is distorted by badwill and restructuring costs. Table 7 presents the formulated earlier proposal to settle the transaction with adjustments of EVA concerning bedwill and restructuring costs.

Table 7. Investment evaluation PV of EVA with badwill, cost of restructuring and EVA adjustment

Year of forcast	0	1	2	3	4	5	…
EVA for purchaser		1,0	1,4	1,2	1,3	0,6	0,8
NOPAT (49%)		1,5	1,9	1,8	1,7	1,2	1,2
Cost of restructuring		-1,5	-1,5	-1,5			
Badwill		4,8					
Cost of Capital		-0,5	-0,5	-0,6	-0,7	-0,6	-0,5
IC		5,0	4,8	5,8	6,8	6,1	4,7
EVA adjutment		-3,3	1,5	1,5	0,3		
IC		5,0	4,8	5,8	6,8	6,1	4,7
IC 49%		9,8	8,1	7,6	7,1	6,1	4,7

Goodwill		-4,8	-4,8	-4,8	-4,8		
Cost of restructuring			1,5	3,0	4,5		
Discount Factor	1,00	0,91	0,83	0,75	0,68	0,62	6,21
PV EVA	0,0	0,9	1,2	0,9	0,9	0,4	4,8
Sum of PV(EVA) = NPV	9,0						

In the first year, an adjustment to badwill and restructuring costs was made, in subsequent years restructuring costs were adjusted. Adjustments are recognized in the invested capital. In the fourth year, when the process of restructuring has been completed, badwill and restructuring costs were settled, the difference increased the profit only once. Such an adjustment makes it easier to evaluate the impact of investments on the value for the bidder. At the same time the adjustment does not affect the evaluation of the investment, the NPV of the adjustment amounts to zero.

4. CONCLUSIONS

Value measures, including Economic Value Added can support investment decision-making, especially in the area of merges and acquisitions. They can constitute an important element of value controlling in companies using VBM systems. The conducted discussion and analysis of the case study indicate that:

1. The presented approach does not differ significantly at the level of the obtained results from the classical income approach based on the NPV. However, the presentation of project valuation using VBM measures carries significant added information on the value. It allows to make independent evaluation of the target company's ability to build value and evaluate it from the investor's point of view.
2. The amount of invested capital does not influence the valuation results when the DCF method in relation to EVA is used.
3. The accounting treatment of goodwill and badwill resulting from the acquisition transaction is crucial to appraising the investment, especially if the investment constitutes a part of the company's shares (not 100%). The complexity of accounting law causes that use of VBM measures requires their additional adjustments while evaluating investment projects.
4. In the case of transactions where the shares price exceeds the value of the acquired assets, the evaluation from the investor's point of view should include part of the EVA generated by the target company and the entire value of the goodwill. As a result, the investor's invested capital equals the price of shares.
5. In the case of transactions where the price of shares exceeds the value of the acquired assets, in order to better evaluate the investment from the VBM measures' point of view, restructuring costs should be matched with badwill. The allocation

method cannot influence evaluation of the transaction effectiveness, however, allows to fully investigate economic sense.

REFERENCES

[1] ANAND, M., GARG, A., ARORA, A. Economic Value Added: Business Performance Measure of Shareholder Value, Management Accountant, 1999, 34 (5), pp. 351–356
[2] BURKSAITIENE, D. Measurement of value creation: Economic Value Added and Net Present Value, Economics & Management. 2009, p709-714. 6p. 2
[3] D'ALAURO, G., The impact of IAS 36 on goodwill disclosure: Evidence of the write-offs and performance effects.. Intangible Capital. 2013, Vol. 9 Issue 3, p754–799. 46p.
[4] DILON, R. D.; OWERS, J. E., EVA as a Financial Metric: Attributes, Utilization, and Relationship to NPV, Financial Practice & Education. Spring/Summer97, Vol. 7 Issue 1, p32-40
[5] FERNANDEZ P., Valuation and Common Sense, 3rd edition, University of Navarra - IESE Business School, 2013
[6] HILL Charles W. L. and GARETH R., Essentials of Strategic Management, Jones 2011
[7] MAGNI, C., A., CAPM and Capital Budgeting: Present/Future, Equilibrium/Disequilibrium, Decision/Valuation, IUP Journal of Financial Economics. Mar 2010, Vol. 8 Issue 1/2, p. 7–35. 29p. 8 Charts
[8] RAȚIU V., M, TUDOR R., V., TIRON, A., The classification of goodwill- an essential accounting analysis.. Review of Economic Studies & Research. Dec2013, Vol. 6 Issue 2, p137–149. 13p. ,
[9] O'BYRNE S., F. , YOUNG D., S., Why Capital Efficiency Measures Are Rarely Used in Incentive Plans, Journal of Applied Corporate Finance, Volume 21 , Number 2 ,Spring 2009
[10] YOUNG, S. D., Some Reflections on Accounting Adjustments and Economic Value Added. Journal of Financial Statement Analysis. Winter 1999, Vol. 4 Issue 2, p7. 13p
[11] WORTHINGTON, A., C. & WEST T., A. Economic Value-Added: A Review of the Theoretical and Empirical Literature. Asian Review of Accounting 2001 9(1), pp. 67–86

Capital structure, cost of capital, dynamic mode,
probability distribution, expected value,
risk sensitivity coefficient

Danuta SERETNA –SAŁAMAJ, Agnieszka SZPARA*

ADJUSTMENT OF THE FINANCIAL COMPANY STRUCTURE IN THE PROCESS OF INVESTMENT, INCLUDING THE RISK SENSITIVITY COEFFICIENT

Decisions regarding the management structure of finance in a company have a significant impact on the financial position and consequently the value of the company. Uncertainty surrounding economic activity creates the need for a dynamic view of the structure of finance in a changing environment and especially in the process of investing. Investment decisions inherent risks. Decision theory is a common interest in many different areas of science, including the analysis and support of decision making. The paper presents the dynamic development of the financial structure in the process of investing with regard to sensitivity to risk.

INTRODUCTION

Management of financial structure of the company can be considered as one of the most difficult issues in the whole process of business management. Conscious use of available capital, use of debt provides an opportunity to achieve greater benefits for the owners of the company. Decisions regarding the management of financial structure of a company have a significant impact on the financial position and consequently the value of the company. Uncertainty surrounding economic activity creates the need for a dynamic view of the structure of finance in a changing environment and especially in the process of investing. Investment decisions inherent risks. Decision theory is a common interest in many different areas of science, including the analysis and support of decision making.

* School of Higher Vocational Education in Nysa, Institute of Finance

Typically, the level of risk in the methods of estimating the profitability of investments, is taken into account in the increase in the relevant discount rate.

Instead of the raising the discount rate as the increase in subjectively assessed risk, this paper proposes the use of risk sensitive optimization for forecasting future cash flows. This model should estimate the future cash flows based on the description of the input parameters of the model. If you are using probability distributions sufficiently large number of parameters in the model, there will be described, then the results will be transferred to the calculation of the discounted cash flow to obtain, after proper simulation, information not only on the expected NPV, but also on the likelihood of adoption by the NPV of the respective ranges of values.

1. DYNAMIC MANAGEMENT OF FINANCIAL COMPANY STRUCTURE

The former decades-achievements literature and the results of empirical studies did not allow to construct a universal optimal financial structure of company or even an approximate model for the construction of [1,2,4,5,6,7,9,10]. Thus, despite some obvious observations regarding the accuracy of the deterministic nature of the formation of the cost of capital, its management structure under real conditions is very difficult. In many cases, it is still almost entirely intuitive and is based on the objective other than to minimize the cost of capital

Since the value of the company calculated in accordance with the DCF model in addition to the cost of capital as the discount rate takes into account the cash flow, attempts to assess whether cash flows can have a random character. Analysis of distributions of net cash flow indicates the possible randomness of changes in these variables.

The concept of dynamic management structure of corporate finance is the process of adapting the structure of finance to a level at which the value of the company will be the highest, at the lowest cost of capital and the level of risk at an acceptable level consistent with the policy and strategy of the company. Dynamic management of financial structure will enable to monitor effectively the financial situation of the company, and to maximize the value of the acceptable level of risk.

The proper way to take into account the risks associated with the uncertainty of future cash flows is to study their distribution, depending on the variety of factors affecting future cash flows. In spite of raising the cost of equity as subjectively assessed risk increase it is proposed to build a financial model described by the parameters of a stochastic distribution. If we are using

probability distributions, we will describe a sufficiently large number of parameters in the model, the results will move to the calculation of the discounted cash flow, after proper simulation we will get information, not only about the expected value of the company, but also about the likelihood of the adoption of appropriate ranges of values.

The proposed methodology is to determine the impact of factors on the two categories (FCF and the discount rate - the cost of capital) occurring in calculating the value of the company using the DCF method. The starting point for the dynamic management of the structure of finance is to identify the stage 0. It should identify the most critical factors affecting the financing structure, and thus equity and debt. The given factors should be divided into determined, random and stochastic.

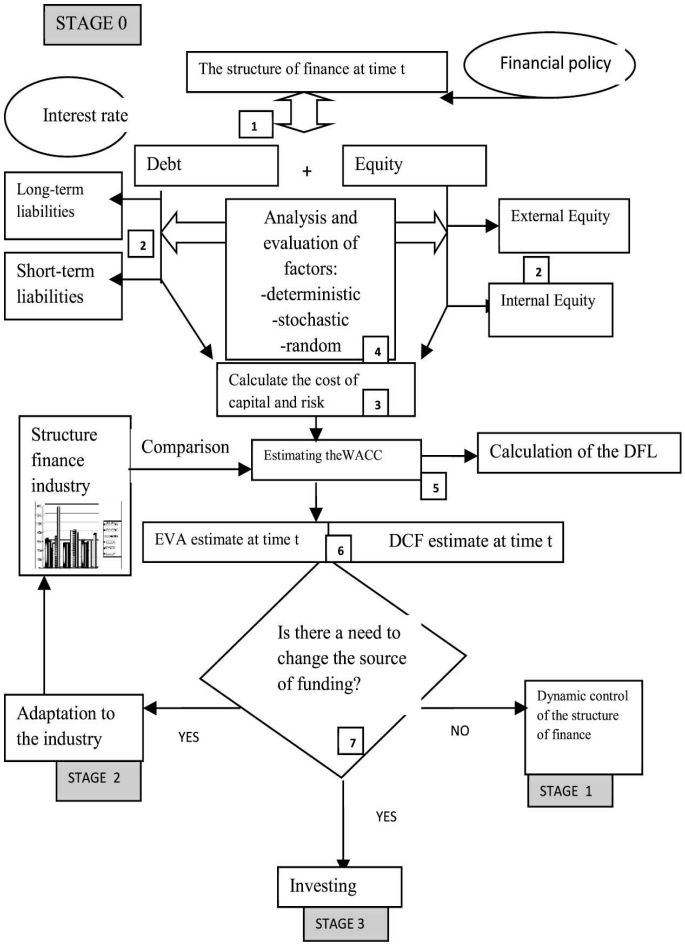

Fig. 1. Methodology dynamic management corporate structure of finance at the stage 0
Source: Own study.

The methodology of the case is divided into four possible phases [0, 1, 2, 3].

Stage [0] refers to the current situation of the company and is based on the estimation of the debt structure, factors affecting it and the calculation of the WACC. These actions aim at calculating the value of the company using the DCF method (Fig. 1).

Procedure for the dynamic management of the financial structure at stage [0] requires analyzing the following blocks:

Block 1: Determination of the current level of use of debt and equity.

Block 2 Determination of parameters associated with each source of financing for equity and debt.

Block 3 Estimation of the capital cost used in financing companies and the risks associated.

Block 4 Estimation of the factors affecting the financial structure of the company.

Block 5 DFL calculation and comparison of WACC for the industry.

Block 6 Estimation of EVA and DCF. After calculating the cost of capital financing the company's assets, it's value is calculated using the DCF method. In DCF, it is necessary to forecast CF, which may include a random factor

Block 7 Decisions about changing the financial structure of a company. If for example, previous studies have shown the risk level incompatible with the financial strategy, too high cost of capital, you should try to match the structure of the branch average [Stage 2]. Satisfactory result requires a dynamic control of financial structure[Stage 1]. Companies wishing to enlarge the status of existing assets should follow the procedure [Stage 3].

Investing in a company depends on capital, which the company has and the ability to obtain additional capital. The cost of using the funding source is the discount rate in the NPV. The method of financing investment in addition to the impact on the cost of capital also affects the cash flow generated that both belons to owners and creditors (Fig. 2).

The basis of methods for estimating the profitability of investment projects are developed by them CF_t cash flows. Forecasting cash flows, we rely on econometric or stochastic methods depending on the occurrence of time-dependent randomness. In NPV method cash flows are discounted and respectively compared with capital expenditures. WACC is assumed as the discount rate. The cost of capital is dependent on the method of financing. Simulation NPV deal with forecasting CF and possible combinations of WACC. Simulation results are the basis for elimination from further consideration of projects with NPV <0, unprofitable projects.

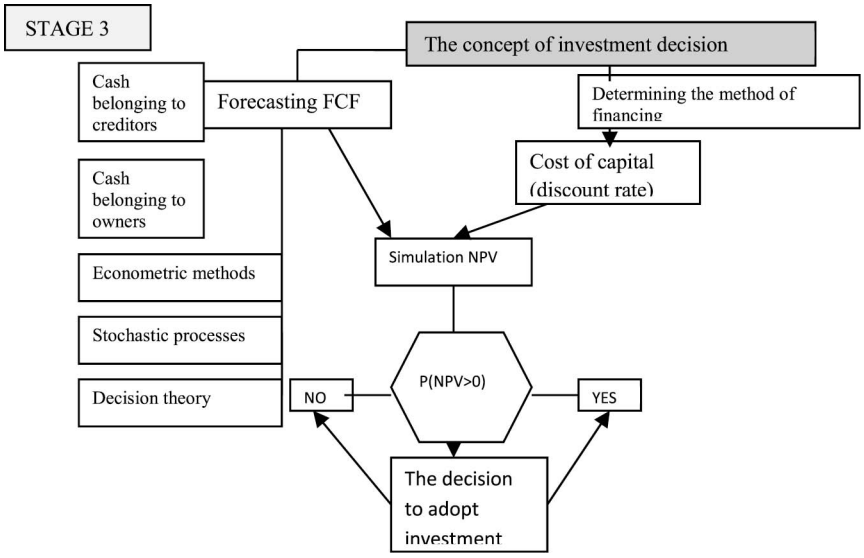

Fig. 2 Methodology dynamically adapt the structure of corporate finance in the investment process
Source: Own study.

2. OPTIMIZATION WITH REGARD TO RISK SENSITIVITY

Let us consider the problem of choosing an alternative a from a set A, whose result will be a random payoff X_a with a-dependent probability distribution Decision-maker, which is insensitive to risk will choose a for which the expected value EX_a is a maximum. To choose a good investment portfolio Markowitz (see example [3,8]), suggested that the optimization criterion be based on maximizing the sum

$$M_a = EX_a + \frac{r}{2} Var(X_a), \tag{1}$$

where r<0 is a decision-maker's risk sensitivity coefficient. Taking this formula into account it is assumed that the decision maker cares about the highest possible expected value as high as possible and the smallest possible variance of the random payment. Somewhat broader and more subtle approach to this problem has stems from utility theory and stochastic optimization (see [3,8,11]) and is to maximize

$$W_a = \frac{1}{r}\ln(Ee^{rX_a}), \tag{2}$$

Using the Taylor expansion for exponential function e^x and for logarithmic function $ln(1+y)$ we can prove, that W_a is the limit of a series containing $\mathrm{E}X_a$ and further the central moments of the random variable X_a multiplied by the number dependent on the powers of *r* converge rapidly to zero. When *r* <0 is close to zero, we have

$$M_a \approx EX_a + \frac{r}{2}Var(X_a). \tag{3}$$

In the case where the random variable X_a has a normal distribution in the above formula have equality. For other distributions equality does not exist, but you can see that the proximity factor *r*<0 to the number zero (talking about a relatively moderate sensitivity risk) means that the selection criterion based on the form of the W_a is close to the idea of Markowitz.

Let us note that when r<0 ,maximization of W_a is equivalent to minimizing $ln(Ee^{rX})$, and because the logarithm is a decreasing function, it is equivalent to minimizing the amount of Ee^{rX_a}.

For illustration, we give a simple example.

EXAMPLE 1

Let $A = \{1,2\}$. Let us assume that X_1 has a value of -1, 1 with equal probabilities. Let $X_2 = 2X_1$. Intuition tells us that careful decision-maker will prefer X_2due to the smaller variance. Both random variables have zero expected values and are indistinguishable by a person insensitive to risk. Let us note that

$$Ee^{rX_1} = \frac{e^{-r} + e^{r}}{2} > Ee^{rX_2} = \frac{e^{-2r} + e^{2r}}{2}, \tag{4}$$

when $r < 0$. For example, for r =-0,1 we have $Ee^{rX_1} \approx 1{,}005 < Ee^{rX_2} \approx 1{,}02$ that is, the above-mentioned criterion is consistent with our intuition. Of course, it can be decisive in situations where intuition is impossible due to more complex random variables and their probability distributions.

In a similar way we can compare the random variables *X* and *Y*. We say that *X* is better than *Y*, where

$$\frac{1}{r} ln(Ee^{rX}) > \frac{1}{r} ln(Ee^{rY}), \tag{5}$$

where r <0 is a fixed factor of risk sensitivity. Obviously, the above inequality is equivalent to the inequality: $Ee^{rX} < Ee^{rY}$.

Let us assume that X is a random result of certain investments. We can assume, for example, that X = NPV. The decision-maker insensitive to the risk will say that investment is acceptable if EX ≥ 0. On the other hand, a person sensitive to risk (characterized by a factor of risk sensitivity r <0) will tell you that the investment is acceptable if

$$\frac{1}{r} ln(Ee^{rX}) > \frac{1}{r} ln(Ee^{0}) = 0 \tag{6}$$

or equivalently

$$Ee^{rX} \leq e^{0} = 1 \tag{7}$$

In the next part we will apply the criterion (7).

3. APPLICATION FOR ASSESSMENT OF INVESTMENT WITH RANDOM CASH FLOWS

Let us consider the investment, in which the initial contribution is I, and the income will be received at the end of each period $t = 1, \dots, n$. It is assumed that these income are random variables $C_1, C_2, \dots, C_n$. The random variable $Z = NPV$ has a known formula:

$$Z = \sum_{t=1}^{n} \frac{C_t}{(1+d)^t} - I \tag{8}$$

where d > 0 is the discount rate per period. Let us designate: $b_t = \frac{1}{(1+d)^t}$; $t = 1, \dots, n$. Then we can simply write that

$$Z = NPV = b_1 C_1 + b_2 C_2 + \cdots + b_n C_n - I \tag{9}$$

Using the criterion (7), we say that the investment is acceptable if the

$$Ee^{rZ} \leq e^0 = 1 \tag{10}$$

Let us suppose that the random variables $C_1, C_2, \dots, C_n$ are independent. From the theory of probability, we know that the random variables $e^{rb_1C_1}, e^{rb_2C_2}, \dots, e^{rb_nC_n}$ are also independent and therefore

$$Ee^{rZ} = E\left(e^{r(b_1C_1+b_2C_2+\cdots+b_nC_n)}e^{-rI}\right) = e^{-rI}\prod_{t=1}^{n} Ee^{rb_tC_t} \tag{11}$$

Thus, the investment is acceptable if the

$$\prod_{t=1}^{n} Ee^{rb_tC_t} \leq e^{rI} \tag{12}$$

Suppose further that the random variables C_t have the same probability distribution. Then we can write (12) in a more transparent formula.

2.1 Let us assume that each random variable has a normal distribution N($\mu;\sigma$).. Then $\mu = EC_t$ and $\sigma^2 = Var(C_t), t = 1, \dots, n$ From formula (17) in the Appendix (see below), we know that if Y has the distribution $N(\mu_t;\sigma_t)$ is

$$Ee^{rY=}e^{r\mu_t} + \frac{r^2\sigma_t^2}{2} \tag{13}$$

The variable $Y = b_tC_t$ has a distribution $N(b\mu_t; b\sigma_t)$. Using the formulas (12) and (13) we obtain

$$\prod_{t=1}^{n} e^{rb_t\mu+\frac{r^2b_t^2\sigma_t^2}{2}} \leq e^{rI}, \tag{14}$$

and hence

$$\sum_{t=1}^{n} rb_t\mu + \frac{r^2b_t^2\sigma_t^2}{2} \leq rI. \tag{15}$$

Dividing both sides of this inequality by r <0 we obtain

$$\sum_{t=1}^{n} b_t \mu + \frac{r b_t^2 \sigma_t^2}{2} \geq I. \tag{16}$$

Substituting $b_t = \frac{1}{(1+d)^t}$; in (16) we obtain conclusions.

CRITERION 1. If cash flows $C_1, \ldots, C_n$ are independent random variables with normal distribution $N(\mu; \sigma)$, the decision-maker characterized by sensitivity to risk at the level of r <0 accept an investment if the

$$\mu \sum_{t=1}^{n} \frac{1}{(1+d)^t} + \frac{r\sigma^2}{2} \sum_{t=1}^{n} \frac{1}{(1+d)^{2t}} \geq I, \tag{17}$$

or

$$\mu \frac{1 - \frac{1}{(1+d)^n}}{d} + \frac{r\sigma^2}{2} \frac{1 - \frac{1}{(1+d)^{2n}}}{d^2 + 2d} \geq I. \tag{18}$$

The condition under Criterion 1 can be generalized to the case where the distributions of the variables C_t are normal $N(\mu_t; \sigma_t)$ but their parameters are different. Derivation of the condition of approval of the project is similar, but it has a less compact form. However, it is relatively simple to apply.

CRITERION 2. We assume independence of random variables $C_1, \ldots, C_n$ and that they have normal distributions. Then the investment project is assumed when

$$\sum_{t=1}^{n} \frac{\mu_t}{(1+d)^t} + \frac{r}{2} \sum_{t=1}^{n} \frac{\sigma_t^2}{(1+d)^{2t}} \geq I \tag{19}$$

CRITERION 3. Cash flows as independent random variables with uniform distribution

Now suppose that C_t has a uniform distribution U [C-c, C+c] on [C-c, C+c], C-c> 0, then

$$E e^{r b_t C_t} = \frac{1}{2c} \int_{C-c}^{C+c} e^{r b_t x} dx = \frac{1}{2c} \left(\frac{e^{r b_t (C+c)} - e^{r b_t (C-c)}}{r b_t} \right) \tag{20}$$

Substituting equation (12) we obtain

$$\prod_{t=1}^{n} \frac{1}{2c}\left(\frac{e^{rb_t(C+c)}-e^{rb_t(C-c)}}{rb_t}\right) \leq e^{rI}, \tag{21}$$

$$\frac{1}{(2cr)^n}\prod_{t=1}^{n}\left(\frac{e^{rb_t(C+c)}-e^{rb_t(C-c)}}{b_t}\right) \leq e^{rI} \tag{22}$$

EXAMPLE 2

Let us consider a three-year investment. We assume that the random variables C_1, C_2, C_3 have uniform distribution on the interval $[900,1100]$. Let us assume that d=0,05.

The decision-maker insensitive to risk can accept the project when

$$1000\left(\frac{1}{1+d}+\left(\frac{1}{1+d}\right)^2+\left(\frac{1}{1+d}\right)^3\right) \geq I$$

$$1000 \cdot \frac{1-\left(\frac{1}{1+d}\right)^3}{d} \geq I$$

$$1000 \cdot \frac{1-\left(\frac{1}{1+0,05}\right)^3}{0,05} \geq I$$

$$I \leq 2723,25$$

Let us consider this example where we have to deal with the decision-maker sensitive to risk. How much will he be able to invest when the risk sensitivity coefficient r = -0.05?

$$\prod_{t=1}^{3}\frac{1}{2 \cdot 100}\left(\frac{e^{-0,05b_t 1100}-e^{-0,05b_t 900}}{-0,05 \cdot b_t}\right) \leq e^{rI}$$

In this case

$$I \leq 2583,23.$$

How much will he be able to invest in other cases?
For the coefficient of r=-0,01

$$I \leq 2683{,}07$$

When r=-0,001

$$I \leq 2719{,}12$$

Noticed difference is relatively small, but the ratio r = -0.001 represents a relatively small sensitivity to risk in the case of amounts in the range of several thousand. The increase in the value of the coefficient of r is increased sensitivity to risk and its effect is visible through the fall in the value of investments and acceptable.

EXAMPLE 3

Let us assume, that each random variable C_t has normal distribution N(1000; 100). Let us consider the 10 years investment. Let us assume that d = 0,05 and r = -0,02.

It is clear that the decision-maker insensitive to the risk will accept an expense I such as

$$\sum_{t=1}^{10} \frac{1000}{(1+0{,}05)^t} \geq I$$

After substituting the data we obtain I ≤ 7721.73. Substituting the data to obtain criterion 1 we will obtain I ≤ 7113.82. You can see that risk awareness somewhat weakens the investment enthusiasm.

APPENDIX

Suppose that $Y \sim N(\mu_t, \sigma_t)$; r<0
Then

$$Ee^{rY} = \int_{-\infty}^{\infty} e^{rx} \frac{1}{\sigma_t\sqrt{2\pi}} e^{\frac{-(x-\mu_t)^2}{2\sigma_t^2}} dx = \frac{1}{\sigma_t\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{rx} e^{\frac{-(x-\mu_t)^2}{2\sigma_t^2}} dx =$$

$$= \frac{1}{\sigma_t\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-(x-\mu_t)^2}{2\sigma_t^2}+rx} dx = \frac{1}{\sigma_t\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-(x-\mu_t)^2}{2\sigma_t^2}+\frac{2rx\sigma_t^2}{2\sigma_t^2}} dx =$$

$$= \frac{1}{\sigma_t\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-x^2+2x\mu_t+2rx{\sigma_t}^2-{\mu_t}^2}{2\sigma_t^2}} dx =$$

$$= \frac{1}{\sigma_t\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-x^2+2x\left(\mu_t+r{\sigma_t}^2\right)-{\mu_t}^2}{2\sigma_t^2}} dx =$$

$$= \frac{1}{\sigma_t\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-\left(x-\left(\mu_t+r{\sigma_t}^2\right)\right)^2}{2\sigma_t^2}+\frac{2\mu_t r{\sigma_t}^2+a^2{\sigma_t}^4}{2\sigma_t^2}} dx =$$

$$= \frac{1}{\sigma_t\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-\left(x-\left(\mu_t+r{\sigma_t}^2\right)\right)^2}{2{\sigma_t}^2}} \cdot e^{\frac{2\mu_t r{\sigma_t}^2+r^2{\sigma_t}^4}{2{\sigma_t}^2}} dx =$$

$$= e^{\frac{2\mu_t r{\sigma_t}^2+r^2{\sigma_t}^4}{2{\sigma_t}^2}} \frac{1}{\sigma_t\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-\left(x-\left(\mu_t+r{\sigma_t}^2\right)\right)^2}{2\sigma_t^2}} dx = e^{\frac{2\mu_t r+r^2{\sigma_t}^2}{2}}$$

Hence

$$Ee^{rY} = e^{\frac{2\mu_t r+r^2{\sigma_t}^2}{2}} = e^{\mu_t r+\frac{r^2{\sigma_t}^2}{2}} \tag{23}$$

SUMMARY

The primary goal of the other objectives is to maximize shareholder value, since it is them, who have the strongest incentive to manage the company in such a way that in the long run, it will have won a competitive battle with other companies. Maximizing the benefit of shareholders, the company maximizes the benefits of other entities associated with it.

The value of the company is affected by the management of a number of factors. These include among others the company's financial structure. Responsible and conscious shaping of the financial structure affects the liquidity generated cash flows, earnings, enterprise value. To achieve and maintain optimal financing structure give benefits the owners, by maximizing the value of the company.

Global changes in economic conditions - increasing volatility of commodity prices, exchange rates, interest rates heighten the risk run by the enterprise business, which enforces the need to consider the impact of risk on the effects of

their decisions on sources of financing. The volatility of the business environment makes it necessary to respond to changes in the environment - companies can adapt to changes in the environment by changing the scale of operations, postponement or abandonment of certain investments, they prove to be unprofitable, etc. Investment decisions making is bound to risk. The risk has an impact on the decision to accept or reject the investment. Decision theory includes the analysis and support of decision-making and can be helpful in determining the sensitivity to risk.

BIBLIOGRAPHY

[1] Benerjee S., Hesmati A., Wilhborg C., The dynamics of capital structure, SSE/EFI Working Paper Series in Economics and Finance, May 2000 nr 333.
[2] Bielawska A., Finanse przedsiębiorstwa. Teoria i praktyka, Wydawnictwo naukowe Uniwersytetu Szczecińskiego, Szczecin 2000.
[3] Bielecki T., Pliska S.R., Economic properties of the risk sensitive criterion for portfolio management. Review of Accounting and Finance 2 (2003), str. 3–17.
[4] Borowiecki R., Czaja J., Jaki A., Strategia gospodarowania kapitałem w przedsiębiorstwie. Zagadnienia wybrane, TNOiK, Warszawa-Kraków 1998.
[5] Brealey R., Myers S.C, Podstawy finansów przedsiębiorstw, Tom1 i Tom 2, PWN Warszawa 1999.
[6] Copeland T., Koller T., Murrin J., Valuation: measuring and managing the value of companies, John Wiley&Sons, New York 1990.
[7] Duliniec A., Struktura i koszt kapitału w przedsiębiorstwie, Wydawnictwo Naukowe PWN, Warszawa 2001.
[8] Fllmer H., Schied A.. Stochastic Finance. Walter de Gruyter, Berlin, 2004.
[9] Modigliani, F., Miller, M., Corporate income taxes and the cost of capital: a correction, American Economic Review 1963
[10] Siudak M., Zarządzanie kapitałem przedsiębiorstwa, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 2001
[11] Whittle P., Risk-Sensitive Optimal Control. Wiley, New York,1990.

Portfolio selection problem, multi-criteria decision making problem, 2-additive measure

Maciej ZARĘBA*

APPLICATION OF 2-ADDITIVE MEASURE TO OPTIMIZATION OF DIVERSIFICATION ACROSS SECTORS IN PORTFOLIO SELECTION PROBLEM

In this paper portfolio selection problem is considered as a multiple criteria decision making problem. The proposed method is based on two stages: the first step is to select a set of shares that meet two basic criteria: are characterized by low risk per unit of expected return on investment and by a low book value to market value ratio. Based on this limited set of securities, in second step model use a Choquet integral to choose a portfolio with maximum diversification across sectors. The model is tested on a data from Warsaw Stock Exchange. The main aim of this paper is to shows how to use Choquet integral to maximize diversification among industries.

1.1. FINANCIAL PORTFOLIO SELECTION PROBLEM

There is a number of reasons why Multi-criteria Decision Making (MCDM) tools are well suited for modeling a capital assets portfolio selection problem [18]. One of the most obvious is that on the market coexist many investors who are differ in many aspects. The level of risk aversion, the size of investment capital, knowledge, experience in investing, investment objectives are only part of the dimensions in which investors differ between each other. New Media Age and increasing amount of associated with the capital market information are affecting positive on reducing inequalities in access to information on the one hand. But on the other hand investors can expand the number of criteria that they use to selection of the shares to the portfolio. In such circumstances, it is reasonable to consider the portfolio selection problem as a multi-criteria decision making problem.

The main aim of portfolio selection is to optimize the allocation of investor's wealth across an investment opportunities. In this paper we deliberately narrowed the set of investment opportunities only to shares and every time when we mention in this

* Wroclaw University of Technology, Institute of Organisation and Management

article about the portfolio selection problem (or financial portfolio selection problem) we have in mind a portfolio of common stocks.

1.2. SELECTION OF CRITERIA

Nowadays, when Financial Crises shook of well-known investment and portfolio theories we should reflect from the beginning on what criteria use on different markets and branches. There is no consensus in the literature what is the best set of criteria for assessing whether particular companies or entire portfolios. Many authors disagree in terms of the type of criteria as well as in terms of the optimal number of criteria [18]. In this article we will focus on three basic criteria:

- Book Value to Market Value Ratio (BV/MV) – companies with a low value of this indicator are usually suspected of overvaluation. However it is possible that a company with a high value of this ratio is still undervaluated[1]. But recent studies on the use of this indicator suggests that even portfolios build by methods based only on this one criterion (BV/MV) are able to significant exceed market benchmarks [9,17].
- Coefficient of variation (CV) – also known as a unitized risk or relative standard deviation is defined as the ratio of the standard deviation of historical share price to the mean of historical values of this price. This ratio can be interpreted as the amount of risk per unit of excpected return and generally the lower value of this ratio the better [17].
- Diversification level among industries which take into account not only number of industries but also dependencies between them (detail information about method using to deal with modeling these dependencies are available in the section describing the model).

2. FINANCIAL PORTFOLIO SELECTION PROBLEM AS A MULTICRITERIA DECISION MAKING PROBLEM

A multicriteria decision making problem can formally be defined as a triple $(A, \mathcal{F}, (\succcurlyeq_k))$ where [6,10,18]:
$A = \{a_1, a_2, \dots, a_n\}$ is a finite set of *alternatives* (e.g. *securities)*
$\mathcal{F} = \{f_1, f_2, \dots, f_m\}$ is a finite set of *criteria* with a finite set of possible values for each of them;

[1] Good example of this situation is a company in which a large share of the assets is intellectual capital, goodwill, etc.

$\succcurlyeq_k$, $k = 1, \ldots m$, is a *preference relation* (a weak order) over A under criterion f_k.

In this work A is a set of portfolios, F is a set of m industries that possible values of its elements is between 0 (in the portfolio is no stock from sector with that weight) and 1 (the capital is invested in the company from one industry).

The values of each criteria should reflect the preference relation, i.e.:

$$f_k(a_i) \geq f_k(a_j) \Leftrightarrow a_i \succcurlyeq_k a_j \tag{1}$$

Selection of the method of preferences aggregation (from the "local" preferences $\succcurlyeq_k$ into one global preference $\succcurlyeq$) is one of the most difficult task in multicriteria decision making problem. That aggregation operator should reflect the actual preferences of the investor with respect to the all investment opportunities and all criteria, i.e.:

$$F(f_1(a_i), \ldots, f_m(a_i)) \geq F(f_1(a_j), \ldots, f_m(a_j)) \Leftrightarrow a_i \succcurlyeq a_j \tag{2}$$

Where: F is an *aggregation operator* that returns an aggregated numerical value for the vector of the evaluations of a given alternative a_i.

2.1. CHOQUET INTEGRAL AS A AGGREGATION OPERATOR

The Choquet integral has been introduced by Choquet [19] in 1953 and now – after years of intense studies also in the field of multicriteria decision analysis [1,2,3,4,7,10, 11, 12, 13, 14] - is one of the most powerful tools in multicriteria decision making. The greatest advantages of this tool is that it allows us to model the interactions between criteria. Simply it is an integral on a fuzzy measure (non-additive measure) and can model many different aggregation operators such as: the weighted sum, the minimum, the maximum, ordered weighted averaging aggregation operator (OWA), weighted ordered weighted aggregation operator (WOWA)[2], etc. [5]. Because of the limitations associated with the size of this work, we present only the basis definition of fuzzy measures and integrals which are necessary for the understanding of the paper, more details of this concept can be find in .

The main foundation of the Choquet integral is the fuzzy measure theory. Let X be a finite set and $P(X)$ be the power set of X, then:

Definition 1 [4]*A function μ: $P(X) \rightarrow [0, \infty)$ is a σ-additive measure if the following axioms are satisfied:*

---

[2] Main disadvantage of using aggregation operators based on additive measure (e.g. weighted sum) is that they can provide only set of potential optimal solutions which is the convex envelop of the feasible solutions set [10].

(1) $\mu(\emptyset) = 0$,
(2) if $A_i, i = 1,2,\ldots,n$, are disjoint subsets of X then $\mu(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i)$.

Definition 2 [4] *A fuzzy measure μ on X is a function μ: P(X) → [0, 1], satisfying the following axioms:*
(i) $\mu(\emptyset) = 0$,
(ii) $A \subseteq B \subseteq X$ implies $\mu(A) \leq \mu(B)$.

Notice that in opposite to a traditional additive measures in fuzzy measure the additivity property has been replaced with a weaker axiom of a monotonicity – property (ii). But models based on this approach need a wider set of parameters[3], therefore, to find a compromise between the number of needed parameters and the expressiveness the k-additivity measure will be introduce.

Definition 3 [14] *A fuzzy measure μ on X is said to be k-order additive it its corresponding pseudo-Boolean function is a multilinear polynomial of degree k, i.e. $a_T=0$ forr all T such that $|T| > k$, and there exist at least one subset T of k elements such that $a_T \neq 0$.*

Its implies that a 2-order additive fuzzy measure can be determined solely by the coefficients μ_i *and* μ_{ij}. The general formula of 2-additive measure is as follow:

$$\mu(K) = \sum_{\{i,j\} \subset K} \mu_{ij} - (|K| - 2) \sum_{i \in K} \mu_i \tag{3}$$

The next step is to use Choquet Integral to aggregate the preferences:

$$C_\mu(a_j) := \sum_{i=1}^{m} \left(f_{(i)}(a_j) - f_{(i-1)}(a_j)\right) \mu(\mathcal{F}_{(i)}), \tag{4}$$

where $0 \leq f_{(1)}(a_j) \leq \cdots \leq f_{(m)}(a_j)$, $\mathcal{F}_{(i)} := \{f_{(i)}, \ldots, f_{(m)}\}$, and $f_{(0)}(a_j) = 0$. Notice that when fuzzy measure is additive, then Choquet Integral takes the form of a simple weighted average:

$$C_\mu(a_j) := \sum_{i=1}^{m} \mu(\{f_i\}) f_i(a_j), \tag{5}$$

[3] For m criteria, one needs 2^m-2 weights to use a Choquet integral based on a capacity (fuzzy measure), where only p-1 weights are needed to use a weighted sum.

2.2. MODEL OF POTFOLIO CONSTRUCTION

On the picture 1 the diagram of the model is presented:

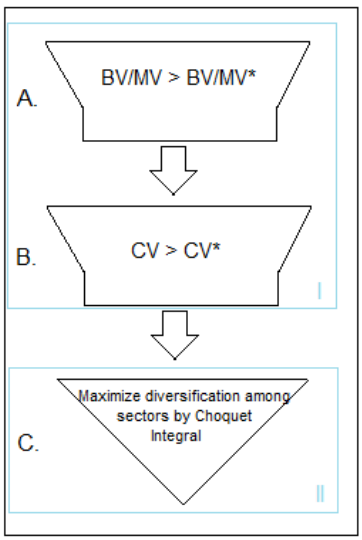

Fig.1. Diagram of the model
Source: own work

First part (module I) consist of two steps A and B, at the entrance of the module I (these are the input data of the whole model as well) are all shares available to the investor: $A = \{a_1, a_2, \dots, a_n\}$, but it could be easy extended to the all investment opportunities (securities, commodities, etc.). Steps A and B eliminate from the set of alternatives these shares which do not meet two necessary conditions:

$$\text{A. } f_{\frac{BV}{MV}}(a_j) \leq \frac{BV}{MV}^*$$

Where $\frac{BV}{MV}^*$ is the historical median of values of BV/MV ratio for all analyzed stocks. And $f_{\frac{BV}{MV}}(a_j)$ is the value of this ratio for company which issued j-th investment assets.

$$\text{B. } f_{CV}(a_j^*) \leq CV^*$$

Where a_j^* is the element of A*, and A* is a set of shares that meet condition A. And CV^* is a historical median of values of this ratio for all analyzed stocks as well.

Obviously there could be a different set of criteria and different methods of estimation of benchmark values for them in a first part of model but this is not a most important part of this model and investor can modify this part according to his needs. The selection of such criteria results from the easy availability of data and surprisingly good performance of models based on such measures [9].

Main objective of the second part of the model is to maximize diversification across sectors. Diversification in this case should be understood as a distribution across sectors which takes into account both the probability of growth of the industry (weights of importance) as well as the interactions between these sectors (weights of interdependencies among sectors). For this reason, Choqet Integral based on non-additive measure (fuzzy measure) is well suited as an aggregation operator.

After the initial selection of investment options model provide a set A^{**} of companies that meet criteria A and B. Let's denote the set of possible portfolios constructed of shares from the set A^{**} as a P. Thus, the objective function in the problem of optimizing the diversification should take the following form:

$$C_\mu(p_j) \coloneqq \sum_{i=1}^{m} \left(f_{(i)}^{II}(p_j) - f_{(i-1)}^{II}(p_j)\right) \mu(\mathcal{F}_{(i)}), \tag{6}$$

Where $\mathcal{F}_{(i)}^{II} \coloneqq \{f_{(i)}^{II}, \dots, f_{(i-1)}^{II}\}$ is a permuted set of criteria used in second stages of portfolio selection process, and power of set $\mathcal{F}_{(i)}^{II}$ is equal to number of sectors (or other class in different approach[4]). Value of the function $f_i^{II}(p_j)$ is understood as a share of stock form i-th sector in j-th portfolio, and value of $C_\mu(p_j)$ is an assessment of the quality of diversification across sectors. Generally, the higher value of C_μ the better diversification across sectors and lower investments risk of the portfolio. Due to values of $f_i^{II}(p_j)$ specify the structure of the portfolio, that must meet also the following conditions derived from the portfolio theory [8]:

$$\sum_{i=1}^{m} f_i^{II} = 1 \tag{7}$$

$$f_i^{II} \in [0,1] \tag{8}$$

What means that sum of weights in portfolio cannot be less than zero and greater than 1, and sum of this weights should be equal to 1.

[4] As a set of criteria of distribution could be use also a set of countries, sizes of companies, etc.).

2.3. PRACTICAL IDENTIFICATION OF FUZZY MEASURES

To calculate a fuzzy measure coefficients we need to estimate a weights of importance of sectors and weights of dependencies between them. Sector indices include companies that are part of a WIG (Warsaw Stock Exchange Index) but selected on the basis of sectoral criterion. Therefore, the sectoral index value can be interpreted as a rating of the industry made by investors. These sector indices are based on the methodology of the WIG and include incomes from dividends and salaries rights issues. Unfortunately, only part of the index is trading long enough to be able to make the necessary statistical analysis, so the analysis is conducted on the basis of companies included in the following sub-indices: WIG-BANKI (consisting of companies from the banking sector), WIG-BUDOW (construction companies), WIG-INFO (IT companies), WIG-SPOZYW (food industry companies) and WIG-TELEKOM (telecommunication companies).

The correlation coefficients between these indices and WIG are shows in table 1:

Table.1. Correlation coefficient between sector indices and WIG

	WIG-Banki	WIG-BUD	WIG-INFO	WIG-SPOZYW	WIG-TELKOM
correlation with WIG	0,98	0,73	0,17	0,89	-0,02

Source: own work

Correlation coefficient between WIG and subsector indices can be using to calculate a weight of importance for a securities from a given sector. For example when investor predict the collapse of the stock market then he should give greater weight for WIG-TELEKOM and WIG-INFO than a weight for WIG-BANKI, WIG-BUD or WIG-SPOZYW. And similarly if the investor provide for an increase in the market then should reverse weights: set higher weights for securities from sectors higher correlated with WIG and lower weights for securities form sectors less correlated with WIG). As a good approximation of the weight of a particular sector should also be the beta coefficient of this sector.

According to the Modern Portfolio Theory investor who wants to maximize return on investment at a given level of risk should consider all his assets and liabilities. What's more, he should pay special attention to the interdependence between the components of the portfolio. Because good portfolio is not just a set of individual "good" investments. Further assumption is about the fact that investors are risk aversion, so making a choice between the two types of assets with equal rates of return, they usually choose an investment with less risk. In order to optimize the structure of the portfolio investor should consider as well an interactions between sectors. Correlation coefficient provide good approximation of dependencies among sectors.

Table.2. Correlation coefficients between choosen sector indices.

	WIG-BUD	WIG-INFO	WIG-SPOZYW	WIG-TELEKOM
WIG-BANKI	0,68	0,08	0,88	-0,11
WIG-BUD		0,37	0,61	0,06
WIG-INFO			0,08	0,67
WIG-SPOZYW				-0,05

Source: own work

Table 2 presents matrix of correlation between sector indices. Calculations were based on daily data for the fifteen-year time horizon (1999-2014), and all data are statistically significant[5]. The strongest positive correlations were observed between WIG-BANKI and WIG-SPOZYW indices, while the strongest negative correlation occured between indices: WIG-BANKI and WIG-TELEKOM.

Investor who wants to minimize the risk of an investment should look for such a combination of securities, which are maximally negatively correlated. For example when investor already has a portfolio that is consist mainly of sceurities from WIG-BANKI, then he should consider diversification by buying of scurities from WIG-TELEKOM[6] etc.

To converse the correlation coefficients into a lattice of fuzzy measure coefficients we use a approach from [20], and the general formulas are as follow:

$$\mu(\{f_i^{II}\}) = corr(WIG, f_i^{II}) \tag{9}$$

Where WIG is the market main index value, and f_i^{II} is the value of i-th sector subindex. $corr(WIG, f_i^{II})$ is a correlation coefficient between WIG and f_i^{II}.

$$\mu(\{f_i^{II}, f_j^{II}\}) = \mu(\{f_i^{II}\}) + \mu(\{f_j^{II}\}) - corr(f_i^{II}, f_j^{II}) * \min\{\mu(\{f_i^{II}\}), \mu(\{f_j^{II}\})\} \tag{10}$$

Fuzzy coefficients for subsets with more than two criteria was calculated according to the formula (3).

---

[5] p-value for all calculated data is much lower than the accepted level of significance (a=0,05).

[6] Obviously individual pairs of securities from negatively correlated sectors may be positively correlated with each other, but in general – if the investment is not limited to the purchase of one share of a given sector – diversification across sectors should yield similar results with much less computation than diversification at securities level.

3. EXPERIMENT

The experiment was conducted on quarterly data form the years 2000-2014. Analysis was performed only on stocks that meet the two basic conditions:

(1) Are part of the sector index, which is listed at least since 2000 (WIG-BANKI, WIG-INFO, WIG-BUD, WIG-SPOZYW or WIG-TELKOM)

(2) Are listed on the Warsaw Stock Exchange at least since 2000

The stock price data come from the official website of a Warsaw Stock Exchange (gpwinfosfera.pl) and the financial data of the analyzed companies come from the Notoria service (notoria.pl). All data were divided into two time horizons: training set covering the years from 2000 till 2010, and test set of data from 2011 to 2014. This means that the basis for all statistical ratios (which were used to estimate the parameters of the models) is the training set and all studies of effectiveness were conducted on the test set.

First the correlation between assessment of portfolios and realized rate of return (4 years rate of return) was tested. The results of the proposed n this paper method were compare next with the results of the method of APT[7].

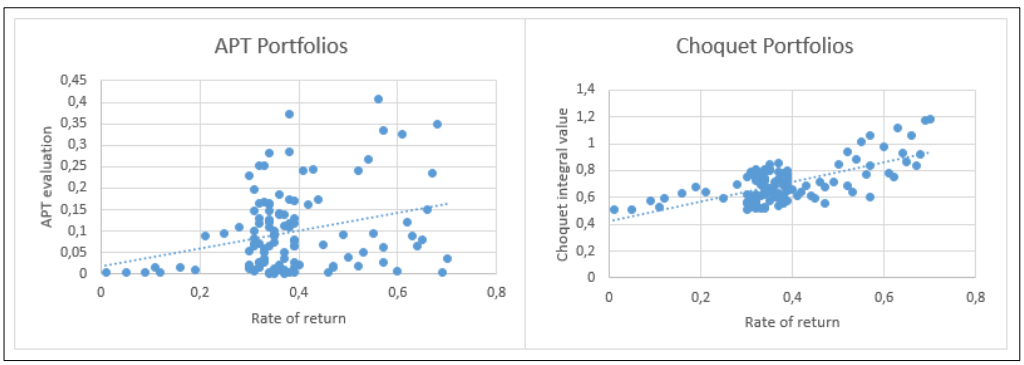

Fig. 2. Comparison of relation between realized rate of return and results of Choquet and APT models. Source: own work

The correlation between rates of return of portfolios and forecasts (evaluations) estimated by Choquet Portfolios for this data is equal 0.69, and for APT model is equal 0.33. It could mean a much higher efficiency of model that is based on the fuzzy measure. Next test was based on a comparison of the results of three portfolios: first

[7] APT assesment for portfolios was made based on parameters calculated with Mean Square Error for k-factors model, more about this method in [15].

build by the method proposed in this work (shortly: Choquet portfolio), APT portfolio and the market portfolio (WIG).

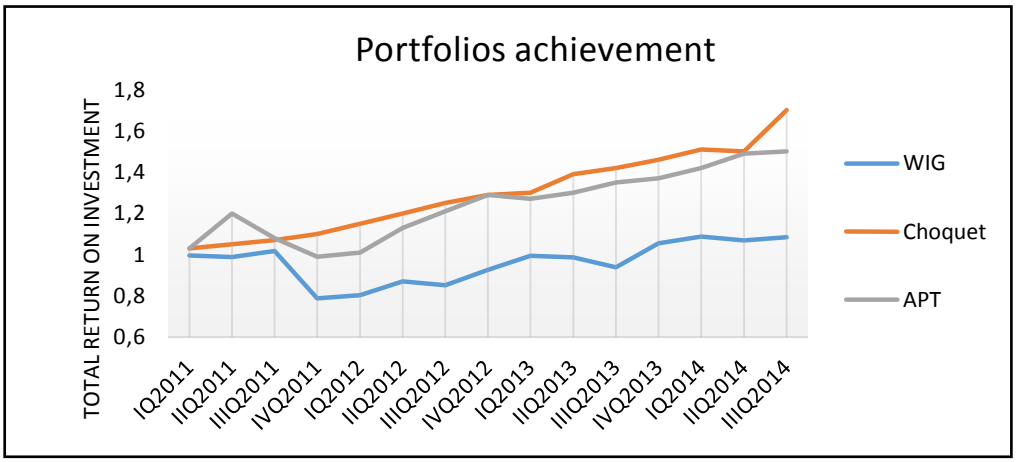

Fig. 3. Comparison of the results of three portfolios
Source: own work

The results of such a survey are presented in figure 3. As You can see APT portfolio and Choquet portfolio far exceeded the benchmark score (WIG), in addition rate of return of Choquet portfolio is characterized by a much lower variability of rate of return than an APT portfolio.

4. CONCLUSIONS

Proposed method of portfolio selection outperforms market benchmark (represent by a rate of growth of WIG) as well as achieve better results than a portfolio build by a one of the APT methods (with additive aggregation operator). The most important advantage of this method is a ability to model an interactions between criteria. As a criteria we can establish many different attributes of the portfolio. In this paper was shown how to use a Choquet integral to determine the structure of weights in portfolio, that allows to maximize the degree of distribution across several sectors. Further development of this model should focus on refining the set of criteria as well as on the use of more sophisticated methods of testing the effectiveness of portfolios.

REFERENCES

[1] GRABISH, M. (1994), *Fuzzy Integral in multicriteria decision making*, Fuzzy Sets and Systems 69.

[2] TAKAHAGI, EIICHIRO,: *An intermediate Evaluation Method among Maximum, Average and Minimum using Fuzzy Measure and Choquet Integral and its Applications*, "Methodologies for the conception, design, and application of intelligent systems; proceedings of the 4th International Conference on Soft Computing 1", World Scientific, 299-302, iizuka, 1996.

[3] TAKAHAGI, EIICHIRO,: *On Identification methods of λ-fuzzy measures using weights and λ, Japanese Journal of Fuzzy Sets and Systems*, vol.12, no.5, 665-676, 2000 (Japanese).

[4] SUGENO, M. (1974): *Theory of fuzzy integrals and its applications*, Doctoral Thesis, Tokyo Insttute of Technology, 1974.

[5] YAGER R.R., *On Ordered Weighted Averaging Aggregation Operators in Multi-criteria Decision Making*, IEEE Transactions on Systems Man and Cybernetics B: 183-190.

[6] ZADEH, L.A., *Fuzzy sets as a basis for a theory of possibility*, Fuzzy Sets and Systems, 1978: 3–28 [7] Grabisch, M., Roubens, M. (1996): Application of the Choquet Integral in Multicriteria Decision Making.

[7] GARMENDIA, L., *The Evolution of the Concept of Fuzzy Measure*, Studies in Computational Intelligence, Springer Verlag, Berlin, 5, 2005: 186-200. [9] Magoc, T., Xiaojing, W., Modave F., Ceberio, M. (2010): Applications of Fuzzy Measures and Intervals in Finance

[8] MARKOWITZ H., *Portfolio Selection*, Journal of Finance 7, 1952, 77-91.

[9] MOSKOWITZ T.J., ASNESS C.S., PEDERSEN L.H., *Value and Momentum Everywhere*, The Journal of Finance, June 2013.

[10] LUST T., ROLLAND A., *2-additive Choquet Optimal Solutions in Multiobjective Optimization Problems*, IPMU 2014, 256-265, 2014.

[11] MORI T., MUROFUSHI, *An analysis of evaluation model using fuzzy measure and the Choquet integral*. In 5th Fuzzy System Symposium, pages 207.212, Kobe, Japan, 1989. In Japanese.

[12] NAKAMORI Y., IWAMOTO, *Analysis of environmental evaluation structure by the Choquet integral model*. In Int. Joint Conf. of the 4th IEEE Int. Conf. on Fuzzy Systems and the 2nd Int. Fuzzy Engineering Symp., pages 695.700, Yokohama, Japan, March 1995.

[13] GRABISCH, M. (1999): The interaction and M¨obius representations of fuzzy measures on finite spaces, k-additive measures: a survey, in: M. Grabisch, T. Murofushi, and M. Sugeno, Eds., Fuzzy Measures and Integrals: Theory and Applications, Springer-Verlag

[14] GRABISCH M., *k-order additive discrete fuzzy measures and their representation*. Fuzzy Sets & Systems 92 (1997), 167-189

[15] ELTON E.J., GRUBER M.J., BROWN J., GOETZMANN W.N., *Modern Portfolio Theory and Investment Analysis*, John Wiley & Sons, 6th Edition, 2003.

[16] UJWARY-GIL A., *Kapitał intelektualny a wartość rynkowa przedsiębiorstwa*. Warszawa, 2009

[17] WILIMOWSKA Z., WILIMOWSKI M., *Sztuka zarządzania finansami (część 1),* Oficyna Wydawnicza Ośrodka Postępu Organizacyjnego, 2001.

[18] TRZASKALIK i in., *Metody wielokryterialne na polskim rynku finansowym*. Polskie Wydawnictwo Ekonomiczne, Warszawa, 2006.

[19] CHOQUET G., *Theory of capacities.* Annales de l'Institut Fourier (5), 131-295, 1953.

[20] KASPERSKI A., ZARĘBA M., *Application of Fuzzy Measures to Financial Portfolio Selection Problem. W: Information Systems Architecture and Technology: Models of Decision Making in the Process of Management in a Risky Environment*. Red. Zofia Wilimowska, L.Borzemski, A.Grzech, J.Świątek. Wrocław, s.55-64, 2013.

Entrepreneurship, high technology,
algorithm, identification

Katarzyna ROSTEK*
Agnieszka SKALA**

IN SEARCH OF HIGH-TECH ENTREPRENEURSHIP: AN ALGORITHM FOR DIFFERENTIATION OF BUSINESS ENTITIES OPERATING WITHIN THE HIGH-TECH SECTOR AS ILLUSTRATED BY THE EXAMPLE OF WARSAW

Manufacturing companies operating within the high-technology sector are of interest to science, industry and national authorities because of the special economic importance attached to them. However, in order to investigate the condition of those companies, support their growth and monitor the effects of the aid awarded to them, it is first necessary to identify the business entities belonging to that sector. A study carried out for this purpose on a population of Warsaw companies has shown that this is not an easy task. Conclusions drawn from the analyses undermine the usefulness of the most commonly applied criterion i.e. the fact of belonging to the industrial sectors considered to be high-tech. To identify the entities belonging to the high-technology industry, it is necessary to perform a sequence of activities which form the procedural algorithm. Usefulness of the algorithm has been verified using the example of a group of Warsaw high-tech companies which were subject to investigation under the European project entitled: "Stołeczne Forum Przedsiębiorczości" (Warsaw Entrepreneurship Forum)[1]. In the future, the algorithm could be used as the basis for the implementation of an IT tool for the identification and description of high-tech businesses.

1. INTRODUCTION

Manufacturing companies operating within the high-technology sector (HT) are of interest to science, industry and national authorities because of the characteristics as-

* Warsaw University of Technology, Faculty of Management, ul. Narbutta 85, 02-524 Warszawa
** Warsaw University of Technology, Faculty of Transport, ul. Koszykowa 75, 00-662 Warszawa
[1] The European Project entitled: "*Stołeczne Forum Przedsiębiorczości* (*Warsaw Entrepreneurship Forum*) – Development, promotion and pilot implementation of new methods of collaboration between the Warsaw authorities and businesses in order to ensure efficient management of economic change", co-funded by the EU under the European Social Fund, agreement no. UDA-POKL.08.01.02-14-137/11.

cribed to them. From the macroeconomic perspective, the sector is important for several reasons. First, its condition is viewed as a reflection of the level of competitiveness and innovativeness in an economy [Acs 2006, Acs et alt. 2012]. As proof of this, it should be noted that the increase in production and particularly in exports of high-tech goods (from 3% of the export value in 2009 to a minimum of 15% in 2030) was mentioned in the government's strategic documents as one of the levers (and simultaneously a metre) of Poland's medium and long-term growth [Polska 2030, p. 70 and 136]. This is in practice equivalent to the allocation of significant public funds to the development of that sector [Polska 2030, p. 77].

Secondly, studies have shown that HT companies are highly productive and competitive, they stimulate the growth of export and increase the level of technology in the entire economy [Zakrzewska-Bielawska, 2011]. Thirdly, high-tech companies create well-paid jobs for the most highly-skilled workforce, which, ultimately, stimulate the domestic demand and encourage savings. An important ingredient of functioning of those companies is their cooperation with science, which contributes to quicker commercialization of scientific research and transfer of technological knowledge [Bromski, 2013].

Moreover, the companies provide an excellent space for research on the latest trends in technologies, management and marketing [Zakrzewska-Bielawska, 2012]. An important consideration here is the accelerating trend towards the reindustrialization of cities, which has also reached the capital of Poland [Westkämper, 2013]. The "purity" of technologies which come under the "high-technology" heading has a beneficial effect on the increase of employment rates among the most highly-skilled workforce, especially among young people. This is a strong case for an active reindustrialization policy with respect to the development of the high-tech sector in metropolitan areas. An incentive here would be the fact that the sector already has a significant presence in Warsaw, which is particularly visible in the central districts of the city [Rostek, Skala, 2014]. All this creates a chance of dynamic development in this direction for other city districts and provides an opportunity for those enterprises which see their opportunity for market expansion in the revival of production (rather than in the services sector only).

In view of the above, it is essential to know which entities are high-tech and which are not. However, an attempt to apply a generally accepted sectoral classification has proved to be ineffective. Use of this classification has resulted in qualifying as high-tech those entities which do not meet the basic classification criteria (e.g. pharmacies instead of drug manufactures, or optical shops instead of companies producing optical equipment) and in omission of important entities which have been assigned to another PKD[2] section, but which are in fact high-tech (e.g. innovative pharmaceutical companies which, due to the high share of trade in medicine in their revenue, are defined as commercial rather than manufacturing companies).

[2] PKD – Polish Classification of Business Activity.

2. FORMULATION OF THE PROBLEM AND RESEARCH OBJECTIVES

Verification of the distinguishing features of HT companies, supporting their development and monitoring the effects of the aid granted to them will only be possible if the manufacturing companies belonging to that sector are correctly identified. Studies carried out for this purpose on a population of Warsaw companies have shown that the task is not as easy as it appears at first glance. The initially identified group of 1363 entities classified as high-tech due to their PKD code denoting the type of business activity was, after a closer examination, reduced to 136 entities (10% of the initial population) which with certainty meet the HT criteria. This is indicative of the seriousness of the problem and its importance for research performed on manufacturing companies. Therefore, the following research questions (RQ) have been formulated:

RQ: What methods should be used in order to ensure a reliable and correct classification of the HT entities?

To identify the entities belonging to the high-tech industry, it is necessary to perform a sequence of activities to verify the basic PKD classification. The activities, arranged in a logical sequence, form the procedural algorithm for verification of criteria and classification of business entities as high-tech. The structure of this algorithm is the objective of this paper (OP):

OP: Preparation of the method for verifying the PKD classification and correct identification of HT entities.

The above objective has been achieved through analysis and discussion of the results of the currently applied HT verification methods (Section 3), proposal of an alternative method (Section 4), verification of its usefulness in a group of Warsaw high-tech companies (Section 5) and discussion of the results obtained (Section 7).

3. METHODS APPLIED IN THE IDENTIFICATION OF HIGH-TECH

High Technology refers to areas of manufacture and products which are highly science-intensive, i.e. are characterized by a high level of R&D intensity [Matusiak 2011]. There exists a commonly applied criterion to distinguish companies operating in the high-tech sector from other companies. The OECD and Eurostat, based on the ratio of research and development expenditure to sales revenues (or value added), distinguish four categories of the manufacturing industry, one of them being high technology [OECD 1005, Eurostat 2014]. The high-tech category comprises three branches: pharmaceuticals, manufacture of electronics, aviation and space industry. Thus, HT companies are identified based on the assumption that all entities performing the listed types of activity qualify as high-tech.

The second criterion is related to the type of manufactured products. Based on this criterion, nine groups of products which qualify as high-tech can be distinguished:

products related to aviation and space industry, computers, electronic and telecommunication products, pharmaceuticals, research and development equipment, electrical machinery, chemicals, non-electrical machinery, weapons and ammunition [Eurostat 2014, GUS 2012]. This criterion is not commonly used due to the lack of statistical data in this area.

The situation is similar for the third criterion which proposes to classify as high-tech those entities which own patents for the products related to the listed categories of products.

In view of the arguments pointing to the lack or restricted access to adequate data, the widest definition of the high-tech industry is applied according to which the HT industry includes all those entities which carry out their activity in the three listed categories: pharmaceuticals, electronics, aviation and aerospace (each of these categories corresponds to specific codes under the PKD classification - Polish Classification of Business Activity).

Following the above guidelines, a study was carried out to identify and analyze the high-tech manufacturing sector located in Warsaw. To this end, the REGON register database (as on December 2012) was obtained from the GUS (National Office of Statistics). The database contained 1363 business entities conducting activity in HT industries together with a list of attributes describing those entities. Even a rough analysis of the obtained database of businesses allowed us to make the following assumptions:

- some of those businesses are probably inactive,
- some of them do not in fact belong to the high-tech sector.

Additional studies were conducted which confirmed the above hypothesis. The research made use of secondary data provided by the report by the PwC consulting firm of September 2011 entitled: "Contribution of the Innovative Pharmaceutical Industry to the Development of Polish Economy" [PwC 2011]. The report contains a detailed analysis of the pharmaceutical market in Poland. A characteristic feature of the market is the clear division into those companies which work to develop new medication (innovative companies) and those which produce the substitutes of the existing drugs (generic companies). According to the report, there are two manufacturing plants producing innovative medication in Warsaw. It has been established that the companies which own those plants have their registered office in Warsaw. However, they were not included in the database of high-tech companies since they defined their main type of activity as commerce rather than manufacturing. This is yet another proof of uselessness of the sectoral classification according to PKD codes.

This brought about a need for developing an alternative method which would enable more accurate and reliable identification of the high-tech sector.

4. HYBRID METHOD FOR HIGH-TECH ENTERPRISE IDENTIFICATION

In order to make the PKD classification more credible, it was deemed necessary to complement the data obtained from the National Office of Statistics (GUS). The complementary information came from two sources: the database run by the ZUS – the Polish National Insurance Company (Business Activity Indicator) and companies' private websites. On this basis, a hybrid method of high-tech enterprise identification was proposed which combines the PKD classification with the company's activity as recorded by the ZUS and the information available online. The high-tech sector identification method which was used to investigate the population of Warsaw companies comprised four steps (Fig.1):

- Step 1 – to extract from the REGON register the entities which are classified as high-tech on the basis of the sectoral criterion (PKD: 21, 26, 30.3);
- Step 2 – to verify the activity of those entities using the Activity Indicator based on the data provided by the ZUS;
- Step 3 – to identify the entity's own website;
- Step 4 – to verify the conformity of the offer published online with the declared type of sectoral activity conducted within the high-tech industry.

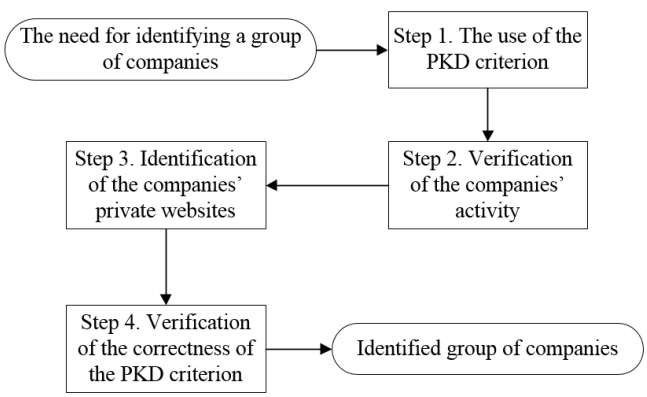

Fig. 1. Hybrid method for high-tech enterprise identification (own compilation)

STEP1.

The analyzed entities were identified based on the main type of activity declared by them which was expressed by the PKD classification code contained in the REGON register. Thus, the first step in the identification process of the population of high-tech companies was to identify those companies which meet the requirement:

$$PKD_i \in (21,26,30.3) \tag{1}$$

where *i* is the *i*-th entity in the database, and 21, 26 and 30.3 are PKD classification codes assigned to three areas of business activity: pharmaceuticals (21), electronics (26), aviation and aerospace (30.30).

Subsequently, out of the set of all attributes in the GUS database those were selected which were regarded as useful classifiers for descriptive analysis. The classifiers are characterized by a finite number of discrete values and informative significance for descriptive and classification analysis. Table 1 shows the list of classifiers within the GUS dataset.

Table 1. Classification attributes within the GUS and ZUS datasets (own compilation)

Item	Name	Description	Type
1	HighTech	PKD class	Characters
2	GD	Municipality identification number	Characters
3	OP	Type of legal personality	Characters
4	FP	Type of legal form	Characters
5	FW	Type of ownership form	Characters
6	FZ_1	Is this a foreign form of ownership?	Numbers
7	LPRAC	Employment size bands	Numbers
8	WWW	Does the company have a website?	Numbers
9	E_I	Export-import activity indicator	Numbers
10	WA	ZUS activity indicator	Characters

STEP 2.

The first step in "purification" process of the database was to eliminate the inactive entities. For this purpose, the Activity Indicator was applied (based on the ZUS database) which limited the list of companies to 558 (41% of the REGON register database). However, additional information had to be obtained in order to conduct further analysis.

STEP 3.

This additional information was obtained from Internet resources where the companies operating in the high-tech sector should publish information on their own websites. To this end, all entities were analyzed and a set of 266 high-tech entities was obtained (48% of all active entities) which were active and had an own website.

STEP 4.

The information obtained from companies' websites was used to verify the conformity of the offer published on the website with the declared type of activity in the GUS database. It was surprising to find that only in 137 cases (51.5% of active entities with an own website) the analysis produced positive results, meaning that a particular company could be classified as actually belonging to the high-tech manufacturing sector.

Based on the above, it can be assumed that classification of a company as high-tech based on the type of activity expressed by the PKD classification code may be subject to substantial error in at least half of the cases (Fig. 2). In light of these findings, unverified (declaratory) PKD classification cannot be a reliable criterion for qualifying a company as high-tech.

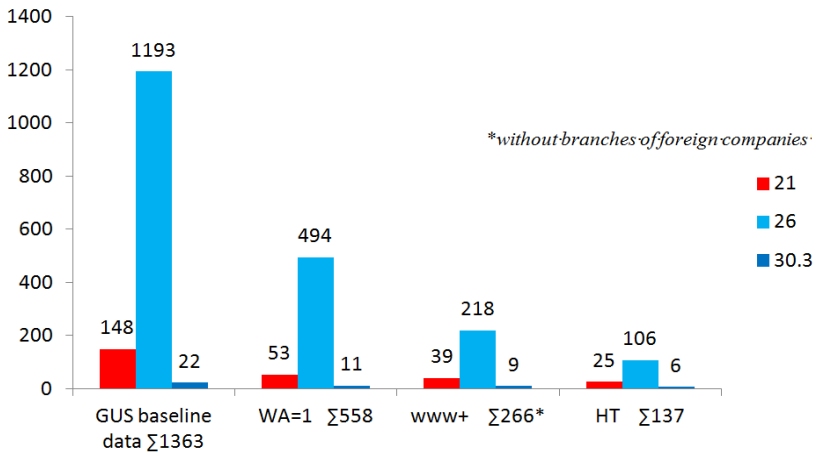

Fig. 2. Structure of the number of companies according to the PKD type of business activity during particular steps of the study (own compilation)

5. VERIFICATION OF THE USEFULNESS OF THE METHOD

The new method increased the reliability of both the identification process and analysis of companies in terms of their high-tech characteristics distinguishing them from other companies.

As regards verification of the correctness of identification, there are examples of companies which were eliminated during particular steps of "purification" of the Warsaw high-tech companies population. In step 2, those companies were eliminated which are not registered as active payers of the mandatory contributions to national insurance (ZUS). Those companies were deemed inactive, and so were the companies which did not employ a single person throughout the year 2012. Step 3 eliminated companies which did not have their own website. It was concluded that the lack of a website is another sign of (market) inactivity of a modern company, especially one operating in the field of advanced technologies. There is of course the risk that such an approach will result in the omission of certain entities. However, in the authors' view, this approach is not affected by a significant error. Step 3 eliminated those entities whose declared PKD was different from the basic type of activity conducted by them.

As regards analysis of the entities in terms of characteristics distinguishing high-tech ventures, it was found that private websites represent a valuable source for enrichment of the data contained in the REGON database. Thus, the next step in the research was to examine the websites of all 137 companies qualified as high-tech in order to provide an in-depth description and identify those characteristics which distinguish the companies from other business entities. The list of those characteristics follows the structure shown in Table 2.

Table 2. Classification attributes in the internet dataset (own compilation)

Item	Name	Description	Type
1	EMPLOY	Does the company employ workers?	Numbers
2	AKTIV	Is the company active? Is the data on the website up-to date?	Numbers
3	SOCIAL	Does the company have social media accounts?	Numbers
4	PATENT	Does the company hold a patent or trademark registration?	Numbers
5	SCIENCE	Does the company cooperate with research centres? Does it have own research laboratories?	Numbers
6	LANG	Does the company have a website in a foreign language?	Numbers
7	B2C	Does the company sell to individual customers?	Numbers
8	B2B	Does the company sell to other companies?	Numbers
9	EXPORT	Does the company export?	Numbers
10	CERT	Does the company have certificates?	Numbers
11	INNOV	Does the company describe itself as innovative?	Numbers

Fig. 3 presents the characteristics of the analyzed population based on the number of companies having particular features.

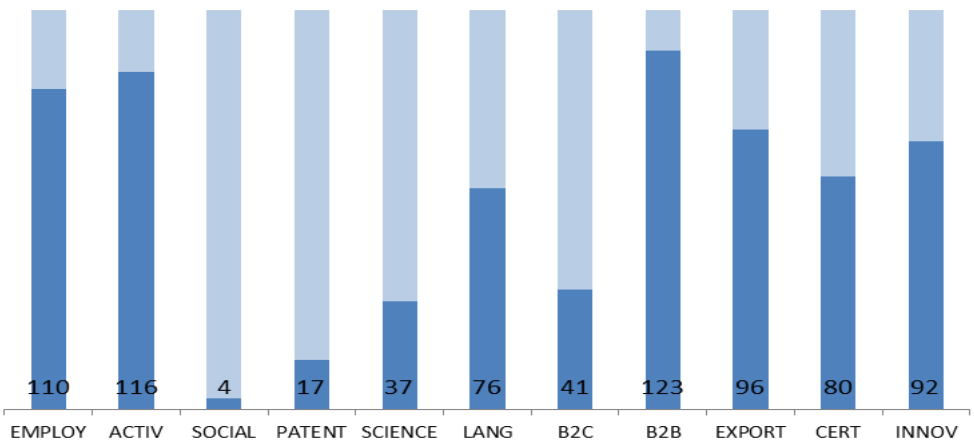

Fig. 3. Number of companies possessing attributes from the internet dataset (own compilation)

An in-depth analysis suggests that high-tech businesses differ from other companies with respect to the following features: conducting business to business (B2B) activity, export activity and searching for new employees. Certain ambiguity arose with respect to some activities related to science: cooperation with research centres, patents and obtaining certification. On the one hand, these seem to be the most important features of the companies associated with high technologies; on the other hand, the number of companies which confirm conducting this type of activity is relatively low. To verify those assumptions, a segmentation analysis of the set was performed using two analytical methods: 1) Ward's clustering method by means of which 3 segments were identified and 2) Kohonen's self-organizing map (SOM) where 4 segments were identified. Segmentation carried out by means of two different methods confirmed that there is considerable variation between the criteria distinguishing particular subgroups of entities within the high-tech population.

Segmentation by means of Ward's method produced a visible gradation of companies in terms of their export activity and scientific activity understood as: patenting, obtaining certification and cooperation in the field of research (or running own laboratories). On the basis of this segmentation, two growth segments were distinguished containing the exporting companies and the scientifically active companies (the segments differ in their level of development). The third segment includes companies which maintain their status quo rather than develop.

Using the self-organizing map, four segments were distinguished. This type of segmentation also distinguished between companies with a lower and a higher level of export and scientific activity. However, this description was enriched by adding new distinctive features. Thus, segment 1 included companies which were active in terms of exports and scientific research, but did not hold patents. Segment 3, on the other hand, comprises companies which are active in terms of exports and hold patents. Segment 4 whose characteristics are very similar to segment 1 has the biggest foreign capital share of all groups. Segment 2, whose specification is similar to that of segment 3 based on cluster analysis, does not compare favourably with other segments. It contains companies which currently do not have many features conducive to development.

In summary of the obtained results, it should be concluded that the strongest criteria differentiating Warsaw high-tech companies are, in the order of importance:

- export activity,
- scientific activity, understood as: obtaining certification, cooperating with research centres (or running own laboratory) or holding patents,
- searching for and employing new workers,
- share of foreign capital in the form of ownership,
- preference to B2B sale over B2C sale.

6. SUMMARY AND DISCUSSION OF THE RESULTS

The studies, whose initial purpose was to provide material for the description of the Warsaw high-tech sector, revealed that there was a major difficulty identifying the entities which qualify as high-tech. As has been found, the basic criterion for qualifying companies as high-tech which is based on the PKD declared by the companies, is not reliable. Therefore, a need arose to define a new method that would provide an unequivocal answer to the question about which companies should and which should not be regarded as high-tech.

Implementation of this method brought two benefits: it verified and restricted the set of entities selected by means of the sectoral criterion (PKD) as well as ensured the enrichment of the analytical dataset. As a result, after subjecting high-tech entities to the four-stage verification process, it was possible to elaborate an in-depth description of those companies, using data obtained from websites. Segmentation performed on the basis of this data demonstrated that the companies considerably differ in their level of development Therefore, further research will focus on finding additional criteria which could be used to identify the high-tech companies with the highest development potential.

In this context, one needs to highlight the significant role of exports which, in the course of the research, proved to be the key feature distinguishing developing business entities. There is a reference here to the "hidden champions" concept (HCh) which has been developed by H. Simon since the mid-1990s [Simon 1996, Simon 2009]. HCh is a special group of businesses which are characterized by: high level of exports, competing on a global scale and a very narrow (and simultaneously profound) specialization of their offer. There are three reasons why those ventures form a valuable part of the economic ecosystem. First, compared to other companies, they are more resistant to negative changes in the economic situation. Secondly, they create new jobs which are characterized by stability and a relatively high level of remuneration. Thirdly, it has been demonstrated that the risk of failure of such companies is much lower than the average level in a given economy. This concept is viewed by the authors as a potential source distinguishing the most promising high-tech businesses and, at the same time, a direction for further research aiming at identifying companies with the highest market potential.

To sum up the findings of the study, we must conclude that the objective set at the beginning of the research project, i.e. elaboration of a reliable and correct method to verify whether a company belongs to the high-tech sector, has been achieved. The method is of a general nature, which means it can also be used to identify other groups and sectors of business entities. Inclusion of the internet data in the presented cycle suggests the need for automation of the process in which the proposed method may provide the basis for designing an IT tool. This will enable further extension of the method by adding elements of quality management of the enriched data, which is essential when using data acquired online.

REFERENCES

[1] Acs, Z., *How is entrepreneurship good for economic growth?*, Innovations, 1(1), 97–107, 2006.
[2] Acs, Z., Audretsch, D., Braunerhjelm, P., Carlsson, B., *Growth and entrepreneurship*, Small Business Economics, 39(2), 289–300, 2012.
[3] Bromski K. (red.), *Współpraca nauki i biznesu. Doświadczenia i dobre praktyki wybranych projektów w ramach Programu Operacyjnego Innowacyjna Gospodarka na lata 2007–2013*, PARP, Warszawa, 2013.
[4] Eurostat, *Eurostat indicators of High-tech industry and knowledge - intensive services. Annex 3 – High-tech aggregation by NACE Rev. 2.*, 2014.
[5] Główny Urząd Statystyczny w Szczecinie, *Nauka i technika w 2011 r.,* Warszawa, GUS, 2012.
[6] Matusiak, K.B. *Wysoka Technika*. w: Matusiak, K.B. (red.), *Innowacje i transfer technologii. Słownik pojęć* (wyd.3, s. 328–331). Polska Agencja Rozwoju Przedsiębiorczości, Warszawa, 2011.
[7] Ministerstwo Administracji i Cyfryzacji, *Polska 2030, Trzecia fala nowoczesności, Długookresowa Strategia Rozwoju Kraju*, Warszawa, 2011.
[8] OECD, *Classification of High-technology Products and Industries*. DSTI/EAS/IND/STP(95)1, 1995.
[9] PwC, *Wkład innowacyjnego przemysłu farmaceutycznego w rozwój polskiej gospodarki, Raport*. 2011.
[10] Rostek K., Skala A., *Warszawskie przedsiębiorstwa wysokiej techniki. Raport analityczny*. Raport niepublikowany w ramach projektu „Stołeczne Forum Przedsiębiorczości", nr umowy UDA-POKL.08.01.02-14-137/11, Warszawa, 2014.
[11] Simon, H., *Hidden champions : lessons from 500 of the world's best unknown companies.* Boston (Mass.): Harvard Business School Press, 1996.
[12] Simon, H., Dietl, M., *Tajemniczy Mistrzowie XXI wieku. Strategie sukcesu nieznanych liderów na światowych rynkach*. Difin, 2009.
[13] Westkämper, E., *Towards the Re-industrialization of Europe: A Concept for Manufacturing for 2030*. Springer Science & Business, Berlin, 2013.
[14] Zakrzewska-Bielawska, A., *Relacje między strategią a strukturą organizacyjną w przedsiębiorstwach sektora wysokich technologii*, Zeszyty Naukowe Politechniki Łódzkiej, Vol. 1095, 2011.
[15] Zakrzewska-Bielawska, A., *The strategic dilemmas of innovative enterprises: proposals for high-technology sectors,* R&D Management, Vol. 42, 4, 2012, 303–314.

BIBLIOTEKA INFORMATYKI SZKÓŁ WYŻSZYCH

Information Systems Architecture and Technology. Advances in Web-Age Information Systems, pod redakcją Leszka BORZEMSKIEGO, Adama GRZECHA, Jerzego ŚWIĄTKA, Zofii WILIMOWSKIEJ, Wrocław 2009

Information Systems Architecture and Technology. Service Oriented Distributed Systems: Concepts and Infrastructure, pod redakcją Adama GRZECHA, Leszka BORZEMSKIEGO, Jerzego ŚWIĄTKA, Zofii WILIMOWSKIEJ, Wrocław 2009

Information Systems Architecture and Technology. Systems Analysis in Decision Aided Problems, pod redakcją Jerzego ŚWIĄTKA, Leszka BORZEMSKIEGO, Adama GRZECHA, Zofii WILIMOWSKIEJ, Wrocław 2009

Information Systems Architecture and Technology. IT Technologies in Knowledge Oriented Management Process, pod redakcją Zofii WILIMOWSKIEJ, Leszka BORZEMSKIEGO, Adama GRZECHA, Jerzego ŚWIĄTKA, Wrocław 2009

Information Systems Architecture and Technology. New Developments in Web-Age Information Systems, pod redakcją Leszka BORZEMSKIEGO, Adama GRZECHA, Jerzego ŚWIĄTKA, Zofii WILIMOWSKIEJ, Wrocław 2010

Information Systems Architecture and Technology. Networks and Networks Services', pod redakcją Adama GRZECHA, Leszka BORZEMSKIEGO, Jerzego ŚWIĄTKA, Zofii WILIMOWSKIEJ, Wrocław 2010

Information Systems Architecture and Technology. System Analysis Approach to the Design, Control and Decision Support, pod redakcją Jerzego ŚWIĄTKA, Leszka BORZEMSKIEGO, Adama GRZECHA, Zofii WILIMOWSKIEJ, Wrocław 2010

Information Systems Architecture and Technology. IT TModels in Management Process, pod redakcją Zofii WILIMOWSKIEJ, Leszka BORZEMSKIEGO, Adama GRZECHA, Jerzego ŚWIĄTKA, Wrocław 2010

Information Systems Architecture and Technology. Web Information Systems Engineering, Knowledge Discovery and Hybrid Computing, pod redakcją Leszka BORZEMSKIEGO, Adama GRZECHA, Jerzego ŚWIĄTKA, Zofii WILIMOWSKIEJ, Wrocław 2011

Information Systems Architecture and Technology. Service Oriented Networked Systems, pod redakcją Adama GRZECHA, Leszka BORZEMSKIEGO, Jerzego ŚWIĄTKA, Zofii WILIMOWSKIEJ, Wrocław 2011

Information Systems Architecture and Technology. System Analysis Approach to the Design, Control and Decision Support, pod redakcją Jerzego ŚWIĄTKA, Leszka BORZEMSKIEGO, Adama GRZECHA, Zofii WILIMOWSKIEJ, Wrocław 2011

Information Systems Architecture and Technology. Information as the Intangible Assets and Company Value Source, pod redakcją Zofii WILIMOWSKIEJ, Leszka BORZEMSKIEGO, Adama GRZECHA, Jerzego ŚWIĄTKA, Wrocław 2011

Information Systems Architecture and Technology. Web Engineering and High-Performance Computing on Complex Environments, pod redakcją Leszka BORZEMSKIEGO, Adama GRZECHA, Jerzego ŚWIĄTKA, Zofii WILIMOWSKIEJ, Wrocław 2012

Information Systems Architecture and Technology. Networks Design and Analysis, pod redakcją Adama GRZECHA, Leszka BORZEMSKIEGO, Jerzego ŚWIĄTKA, Zofii WILIMOWSKIEJ, Wrocław 2012

Information Systems Architecture and Technology. System Analysis Approach to the Design, Control and Decision Support, pod redakcją Jerzego ŚWIĄTKA, Leszka BORZEMSKIEGO, Adama GRZECHA, Zofii WILIMOWSKIEJ, Wrocław 2012

Information Systems Architecture and Technology. The Use of IT Models for Organization Management, pod redakcją Zofii WILIMOWSKIEJ, Leszka BORZEMSKIEGO, Adama GRZECHA, Jerzego ŚWIĄTKA, Wrocław 2012

Information Systems Architecture and Technology. Intelligent Information Systems, Knowledge Discovery, Big Data and High Pefomance Computing, pod redakcją Leszka BORZEMSKIEGO, Adama GRZECHA, Jerzego ŚWIĄTKA, Zofii WILIMOWSKIEJ, Wrocław 2013

Information Systems Architecture and Technology. Networks Architecture and Applications, pod redakcją Adama GRZECHA, Leszka BORZEMSKIEGO, Jerzego ŚWIĄTKA, Zofii WILIMOWSKIEJ, Wrocław 2013

Information Systems Architecture and Technology. Knowledge Based Approach to the Design, Control and Decision Support, pod redakcją Jerzego ŚWIĄTKA, Leszka BORZEMSKIEGO, Adama GRZECHA, Zofii WILIMOWSKIEJ, Wrocław 2013

Information Systems Architecture and Technology. Models of Decision Making in the Process of Management in a Risky Environment, pod redakcją Zofii WILIMOWSKIEJ, Leszka BORZEMSKIEGO, Adama GRZECHA, Jerzego ŚWIĄTKA, Wrocław 2013

Wydawnictwa Politechniki Wrocławskiej
są do nabycia w księgarni
plac Grunwaldzki 13, 50-377 Wrocław
budynek D-1 PWr., tel. 71 320 29 35
Prowadzimy sprzedaż wysyłkową
zamawianie.ksiazek@pwr.wroc.pl

ISBN 978-83-7493-858-7